Group Theory

First Term 2001
D. D. Vvedensky (d.vvedensky@ic.ac.uk)

Office Hours:
Tu 2-3, Fr 11-12 (Blackett 807)

1. **Introduction**
   a. Symmetry in physics
   b. Discrete and continuous symmetries
   c. Symmetry in quantum mechanics

2. **Mathematical Background for Discrete Groups**
   a. Groups
   b. Subgroups
   c. Cosets
   d. Conjugacy classes

3. **Representations of Groups**
   a. Reducible and irreducible representations
   b. Schur's lemmas and the Great Orthogonality Theorem
   c. Character tables
   d. Direct products and their decomposition

4. **Physical Applications of Discrete Groups**
   a. The group of the Hamiltonian
   b. Eigenfunctions and irreducible representations
   c. Bloch's theorem
   d. Selection rules

5. **Continuous Groups, Lie Groups, and Lie Algebras**
   a. Linear transformation groups
   b. Infinitesimal generators
   c. Algebra of infinitesimal generators

6. **Irreducible Representations of SO(2) and SO(3)**
   a. Orthogonality relations and the density function
   b. Basis functions for irreducible representations of SO(2)
   c. Spherical Harmonics and characters for SO(3)

7. **Unitary Groups**
   a. Unitary groups and particle physics
   b. Basis states for SU(N)
   c. Multiparticle states and direct products
   d. Young tableaux

Recommended Books

Basic course material:

H.F. Jones, *Groups, Representations and Physics* (Adam Hilger, Bristol, 1998)
M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw–Hill, 1964)

Related and advanced treatments:

R. Hermann, *Lie Groups for Physicists* (Benjamin, 1966)
D.B. Lichtenberg, *Unitary Symmetry and Elementary Particles* (Academic, 1978)

Chapter 1

Introduction

As far as I can see, all a priori *statements in physics have their origin in symmetry.*
—Hermann Weyl[1]

1.1 Symmetry in Physics

Symmetry is a fundamental human concern, as evidenced by its presence in the artifacts of virtually all cultures. Symmetric objects are aesthetically appealing to the human mind and, in fact, the Greek work *symmetros* was meant originally to convey the notion of "well-proportioned" or "harmonious." This fascination with symmetry first found its rational expression around 400 B.C. in the Platonic solids and continues to this day unabated in many branches of science.

1.1.1 What is a Symmetry?

An object is said to be symmetric, or to have a symmetry, if there is a transformation, such as a rotation or reflection, whereby the object looks the same after the transformation as it did before the transformation. In Fig. 1.3, we show an equilateral triangle, a square, and a circle. The triangle is indistinguishable after rotations of $\frac{1}{3}\pi$ and $\frac{2}{3}\pi$ around its geometric center, or symmetry axis. The square is indistinguishable

[1]In *Symmetry* (Princeton University Press, 1952)

after rotations of $\frac{1}{2}\pi$, π, and $\frac{3}{2}\pi$, and the circle is indistinguishable after *all* rotations around their symmetry axes. These transformations are said to be **symmetry transformations** of the corresponding object, which are said to be **invariant** under such transformations. The more symmetry transformations that an object admits, the more "symmetric" it is said to be. One this basis, the circle is "more symmetric" than the square which, in turn, is more symmetric than the triangle. Another property of the symmetry transformations of the objects in Fig. 1.3 that is central to this course is that those for the triangle and square are *discrete*, i.e., the rotation angles have only discrete values, while those for the circle are *continuous*.

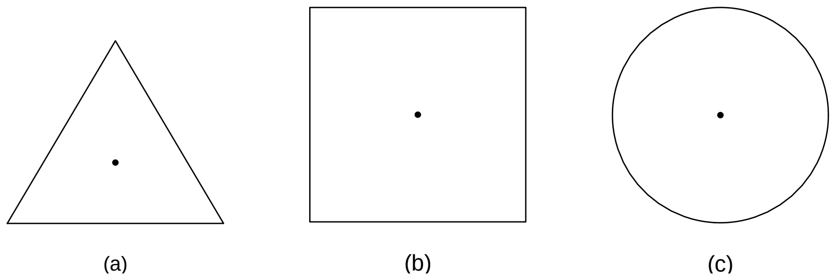


Figure 1.1: An equilateral triangle (a), square (b) and circle (c). These objects are invariant to particular rotations about axes that are perpendicular to their plane and pass through their geometric centers (indicated by dots).

1.1.2 Symmetry in Physical Laws

In the physical sciences, symmetry is of fundamental because there are transformations which leave the laws of physics invariant. Such transformations involve changing the variables within a physical law such that the equations describing the law retain their form when expressed in terms of the new variables. The relationship between symmetry and physical laws began with Newton, whose equations of motion were found to be the same in different frames of reference related by Galilean transformations. Symmetry was also the guiding principle that enabled Lorentz and Poincaré to derive the transformations, now known as Lorentz transformations, which leave Maxwell's equations invariant. The incompatibility between the Lorentz invariance of Maxwell's

equations and the Galilean invariance of Newtonian mechanics was, of course, resolved by Einstein's special theory of relativity.

As an example of a symmetry in a physical law, consider the propagation of an impulse at the speed of light c. This is governed by the wave equation, which is obtained from Maxwell's equations:

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\,. \tag{1.1}$$

The Lorentz transformation of space-time coordinates corresponding to a velocity $\boldsymbol{v} = (v, 0, 0)$ is

$$x' = \gamma(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma\Big(t - \frac{v}{c^2}x\Big)\,, \tag{1.2}$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$. When expressed in terms of the transformed coordinates (x', y', z', t'), the wave equation (1.1) is found to retain its form under this transformation:

$$\frac{1}{c^2}\frac{\partial^2 u'}{\partial t'^2} = \frac{\partial^2 u'}{\partial x'^2} + \frac{\partial^2 u'}{\partial y'^2} + \frac{\partial^2 u'}{\partial z'^2}\,. \tag{1.3}$$

This implies that the wave propagates in the same way with the same velocity in two inertial frames that are in uniform motion with respect to one another. The Lorentz transformation is thus a symmetry transformation of the wave equation (1.1) and this equation is said to be **covariant** with respect to these transformations. In general, symmetry transformations of physical laws involve the space-time coordinates, which are sometimes called **geometrical symmetries**, and/or internal coordinates, such as spin, which are called **internal symmetries**.

1.1.3 Noether's Theorem

Identifying appropriate symmetry transformations is one of the central themes of modern physics since their mathematical expression affects the structure and predictions of physical theories. Work by both mathematicians and physicists, culminating with Emmy Noether, led to the demonstration that there was a deep relationship between symmetry and conservation laws. This is now known as Noether's Theorem:

Noether's Theorem. The covariance of the equations of motion with respect to a continuous transformation with n parameters implies the existence of n quantities, or constants of motion, i.e., **conservation laws.**

In classical mechanics, the conservation of linear momentum results from the translational covariance of Newton's equations of motion, i.e., covariance with respect to transformations of the form $\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{a}$, for any vector $\boldsymbol{a}$. The conservation of angular momentum similarly results from rotational covariance, i.e., covariance with respect rotations in space: $\boldsymbol{r}' = R\boldsymbol{r}$, where R is a 3×3 rotation matrix. Finally, the conservation of energy results from the covariance of Newton's equations to translations in time, i.e., transformations of the form $t' = t + \tau$.

1.1.4 Symmetry and Quantum Mechanics

The advent of quantum mechanics and later quantum field theory fostered entirely new avenues for investigating the consequences of symmetry. London and Weyl introduced a type of transformation known as a gauge transformation into quantum theory, with total electric charge as the conserved quantity. In the early 1960s, Gell–Mann and Ne'eman proposed the unitary symmetry SU(3) for the strong interactions. This led to the proposal by Gell–Mann and Zweig of a new, deeper, level of quanta, "quarks," to account for this symmetry. Heisenberg, Goldstone and Nambu suggested that the ground state (i.e., the vacuum) of relativistic quantum field theory may not have the full global symmetry of the Hamiltonian, and that massless excitations (Goldstone bosons) accompany this "spontaneous symmetry breaking." Higgs and others found that for spontaneously broken gauge symmetries there are no Goldstone bosons, but instead massive vector mesons. This is now known as the Higgs phenomenon and its verification verification has been the subject of extensive experimental effort.

Another aspect of symmetry, also due to the quantum mechanical nature of matter, arises from the arrangement of atoms in molecules and solids. The symmetry of atomic arrangements, whether in a simple diatomic molecule or a complex crystalline material such as a high-temperature superconductor, affects many aspects of their electronic

and vibrational properties and especially their response to external thermal, mechanical, and electromagnetic perturbations. The transformation properties of wavefunctions in quantum mechanics are an example of what is known as **Representation Theory**, which was developed by the mathematicians Frobenius and Schur near the turn of the 20th century. This inspired a huge effort by physicists and chemists to determine the physical consequences of the symmetries of wavefunctions which continues to this day. Notable examples include Bloch's work on wavefunctions in periodic potentials, which forms the basis of the quantum theory of solids, Pauling's work on the chemical interpretation orbital symmetries, and Woodward and Hoffman's work on how the conservation of orbital symmetry determines the course of chemical reactions. Recent scientific advances that highlight the prominent role that symmetry maintains in condensed-matter physics is the discovery of quasicrystals, which have rotational symmetries (e.g., fivefold, as shown in Fig. 1.2) which are incompatible with the translational symmetry of ordinary crystals and are thus sometimes called aperiodic, and the C_{60} form of carbon, known as "Buckminsterfullerene," or "Buckyballs", a name derived from its resemblance to structures (geodesic domes) proposed by R. Buckminster Fuller as an alternative to conventional architecture.

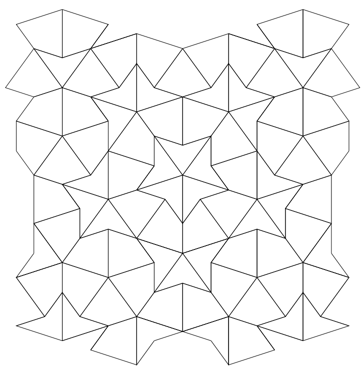


Figure 1.2: A section of a Penrose tile, which has a fivefold rotational symmetry, but no translational symmetry. This two-dimensional structure shares a number of features with quasicrystals.

1.2 Examples from Quantum Mechanics

1.2.1 One-Dimensional Systems

To appreciate how symmetry enters into the description of quantum mechanical systems, we consider the time-independent Schrödinger equation for the one-dimensional motion of a particle of mass m bound by a potential $V(x)$:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+V(x)\right]\varphi(x)=E\varphi(x)\,, \tag{1.4}$$

where $\hbar = h/2\pi$, h is Planck's constant, φ is the wavefunction, and E is the energy eigenvalue. By writing this equation as $\mathcal{H}\varphi = E\varphi$, we identify the coordinate representation of the Hamiltonian operator as

$$\mathcal{H}=-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+V(x)\,. \tag{1.5}$$

In the following discussion, we will utilize the fact that the energy eigenvalues of one-dimensional quantum mechanical problems such as that in (1.4) are nondegenerate, i.e., each energy eigenvalue is associated with one and only one eigenfunction.[2]

Suppose that the potential in (1.4) is an even function of x. The mathematical expression of this fact is the invariance of this potential under the inversion transformation $x \to -x$:

$$V(-x)=V(x)\,. \tag{1.6}$$

Examples of such potentials are the symmetric square well and the harmonic oscillator (Fig. 1.3), but the particular form of the potential is unimportant for this discussion. The kinetic energy term in (1.4) is also invariant under the same inversion transformation as the potential, since

$$\frac{\mathrm{d}^2}{\mathrm{d}(-x)^2}=\frac{\mathrm{d}^2}{\mathrm{d}x^2} \tag{1.7}$$

[2]This follows directly from the fact that this equation, *together with appropriate boundary conditions*, constitute a Sturm–Liouville problem. Other well-known properties of solutions of Schrödinger's equation (real eigenvalues, discrete eigenvalues for bound states, and orthogonality of eigenfunctions) also follow from the Sturm–Liouville theory.

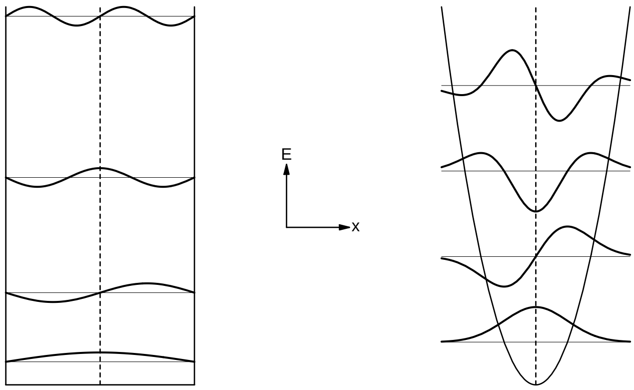


Figure 1.3: The first four eigenfunctions of the Schrödinger equation (1.4) for an infinite square-well potential, $V(x) = 0$ for $|x| \leq L$ and $V(x) \to \infty$ for $|x| > L$ (left), and a harmonic oscillator potential, $V(x) = \frac{1}{2}kx^2$, where k is the spring constant of the oscillator (right). The abscissa is the spatial position x and the ordinate is the energy E, with the vertical displacement of each eigenfunction given by its energy. The origins are indicated by broken lines.

Thus, the Hamiltonian operator in (1.5) is itself invariant under inversion, i.e., inversion is a symmetry transformation of this Hamiltonian. We now use this property of $\mathcal{H}$ to change variables from x to $-x$ in (1.4) and thereby obtain the Schrödinger equation for $\varphi(-x)$:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\varphi(-x) = E\varphi(-x) \tag{1.8}$$

Since E is nondegenerate, there can be only one eigenfunction associated with this eigenvalue, so the $\varphi(-x)$ cannot be linearly independent of $\varphi(x)$. The only possibility is that $\varphi(-x)$ is *proportional* to $\varphi(x)$:

$$\varphi(-x) = A\varphi(x) \tag{1.9}$$

where A is a constant. Changing x to $-x$ in this equation,

$$\varphi(x) = A\varphi(-x) \tag{1.10}$$

and then using (1.9) to replace $\varphi(-x)$, yields

$$\varphi(x) = A^2\varphi(x) \tag{1.11}$$

This requires that $A^2 = 1$, i.e., $A = 1$ or $A = -1$. Combining this result with (1.9) shows that the eigenfunctions φ of (1.4) must be either even

$$\varphi(-x) = \varphi(x) \tag{1.12}$$

or odd

$$\varphi(-x) = -\varphi(x) \tag{1.13}$$

under inversion. As we know from the solutions of Schrödinger's equation for square-well potentials and the harmonic oscillator (Fig. 1.3), both even and odd eigenfunctions are indeed obtained. Thus, not all eigenfunctions have the symmetry of the Hamiltonian, although the ground state usually does.[3] Nevertheless, the symmetry (1.6) does provide a *classification* of the eigenfunctions according to their parity under inversion. This is a completely general result which forms one of the central themes of this course.

1.2.2 Symmetries and Quantum Numbers

The example discussed in the preceding section showed how symmetry enters explicitly into the solution of Schrödinger's equation. In fact, we can build on our discussion in Sec. 1.1.2, and especially Noether's theorem, to establish a general relationship between continuous symmetries and quantum numbers.

Consider the time-dependent Schrödinger equation for a free particle of mass m in one dimension:

$$i\hbar\frac{\partial\varphi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\varphi}{\partial x^2}\,. \tag{1.14}$$

The solutions to this equation are plane waves:

$$\varphi(x,t) = \mathrm{e}^{i(kx-\omega t)}\,, \tag{1.15}$$

where k and ω are related to the momentum and energy by $p = \hbar k$ and $E = \hbar\omega$. In other words, the quantum numbers k and ω of the

[3] A notable exception to this is the phenomenon of *spontaneous symmetry-breaking* discussed in Sec. 1.1, where the symmetry of the equations of motion and the boundary conditions is not present in the observed solution for the ground state.

solutions to Eq. (1.14) correspond to the momentum and energy which, because of the time- and space-translational covariance of this equation, correspond to conserved quantities. Thus, quantum numbers are associated with the symmetries of the system. Similarly, for systems with rotational symmetry, such the hydrogen atom or, indeed, *any* atom, the appropriate quantum numbers are the energy and the angular momentum, the latter producing two quantum numbers, as required by Noether's theorem, because the transformations have two degrees of freedom.

1.2.3 Matrix Elements and Selection Rules

One of the most important uses of symmetry is to identify the matrix elements of an operator which are required to vanish. Continuing with the example in the preceding section, we consider the matrix elements of an operator $\mathcal{H}'$ whose position representation $\mathcal{H}'(x)$ has a definite parity. The matrix elements of this operator are given by

$$\mathcal{H}'_{ij} = \int \varphi_i(x)\mathcal{H}'(x)\varphi_j(x)\,\mathrm{d}x \tag{1.16}$$

where the range of integration is symmetric about the origin. If $\mathcal{H}'$ has even parity, i.e., if $\mathcal{H}'(-x) = \mathcal{H}'(x)$, as in (1.6), then these matrix elements are nonvanishing only if $\varphi_i(x)$ and $\varphi_j(x)$ are both even or both odd, since only in these cases is the integrand an even function of x. This is called a **selection rule**, since the symmetry of $\mathcal{H}'(x)$ determines, or selects, which matrix elements are nonvanishing.

Suppose now that $\mathcal{H}'(x)$ has odd parity, i.e., $\mathcal{H}'(-x) = -\mathcal{H}'(x)$. The matrix elements in (1.16) now vanishes if $\varphi_i(x)$ and $\varphi_j(x)$ are both even or both odd, since these choices render the integrand an odd function of x. In other words, the selection rule now states that only eigenfunctions of *opposite* parity are coupled by such an operator. Notice, however, that the use of symmetry only identifies which matrix elements *must* vanish; it provides no information about the *magnitude* of the nonvanishing matrix elements.

Suppose that

$$\mathcal{H}'(x) = Ax \tag{1.17}$$

where A is a constant, i.e., $\mathcal{H}'(x)$ is proportional to the coordinate x. Such operators arise in the quantum theory of transitions induced by an electromagnetic field.[4] $\mathcal{H}'(x)$ clearly has odd parity, so the matrix elements (1.16) are nonvanishing only if $\varphi_i(x)$ and $\varphi_j(x)$ have *opposite* parity. But, if

$$\mathcal{H}'(x) = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} \tag{1.18}$$

which is the coordinate representation of the kinetic energy operator, then the matrix elements (1.16) are nonvanishing only if $\varphi_i(x)$ and $\varphi_j(x)$ have the *same* parity.

Selection rules are especially useful if there are **broken symmetries**. For example, the Hamiltonian of an atom, which is the sum of the kinetic energies of the electrons and their Coulomb potentials, is invariant under all rotations. But when an atom is placed in an electric or magnetic field, the Hamiltonian acquires an additional term which is *not* invariant under all rotations, since the field now defines a preferred direction. These are the Stark and Zeeman effects, respectively. A similar situation is encountered in quantum field theory when, beginning with a Lagrangian that is invariant under certain symmetry operations, a term is added which does not have this invariance. If the symmetry-breaking terms in these cases are small, then selection rules enter into the perturbative calculation around the solutions of the symmetric theory.

1.3 Summary

The notion of symmetry implicit in all of the examples cited in this chapter is endowed with the algebraic structure of "groups." This is a topic in mathematics that had its beginnings as a formal subject only in the late 19th century. For some time, the only group that was know and whose properties were studied were permutation groups. Cauchy played a major part in developing the theory of permutations, but it was the English mathematician Cayley who first formulated the notion of an abstract group and used this to identify matrices and quaternions

[4]E. Merzbacher, *Quantum Mechanics* 2nd edn. (Wiley, New York, 1970), Ch. 18.

as groups. In a later paper, Cayley showed that every finite group could be represented in terms of permutations, a result that we will prove in this course. The fact that geometric transformations, as discussed in this chapter, and permutations, share the same algebraic structure is part of the richness of the subject and is rooted in its history as an adjunct to the study of algebraic solutions of equations. In the next chapter, we discuss the basic properties of groups that form the basis of this course.

Chapter 2

Elements of Abstract Group Theory

Mathematics is a game played according to certain simple rules with meaningless marks on paper.

—David Hilbert[1]

The importance of symmetry in physics, and for quantum mechanics in particular, was discussed in the preceding chapter. In this chapter, we begin our development of the algebraic structure which enables us to formalize what we mean by "symmetry" by introducing the notion of a group and some related concepts. In the following chapters we will explore the consequences of this algebraic structure for applications to physics.

2.1 Groups: Definitions and Examples

The motivation for introducing an algebraic structure to describe symmetry in physical problems is based on transformations. But the definition of a group is based on a much more abstract notion of what a "transformation" entails. Accordingly, we first set out the conditions

[1]As quoted in, N. Rose, *Mathematical Maxims and Minims* (Rome Press, Raleigh, North Carolina, 1988).

that an *abstract* group must satisfy and then consider both abstract and concrete examples.

Definition. A **group** G is a set of elements $\{a, b, c, \ldots\}$ together with a binary composition law, called *multiplication*, which has the following properties:

1. **Closure.** The composition of any two elements a and b in G, called the *product* and written ab, is itself an element c of G: $ab = c$.

2. **Associativity.** The composition law is associative, i.e., for any elements a, b, and c in G, $(ab)c = a(bc)$.

3. **Identity.** There exists an element, called the **unit** or **identity** and denoted by e, such that $ae = ea = a$ for every element a in G.

4. **Inverses.** Every element a in G has an inverse, denoted by a^{-1}, which is also in G, such that $a^{-1}a = aa^{-1} = e$.

The closure property ensures that the binary composition law does not generate any elements outside of G. Associativity implies that the computation of an n-fold product does not depend on how the elements are grouped together.[2] For example, the product abc is unambiguous because the two interpretations allowed by the existence of a binary composition rule, $(ab)c$ and $a(bc)$, are equal. As will be shown in Sec. 2.3, the left and right identities are equal and unique, as are the left and right inverses of each element. Thus we can replace the existence of an identity and inverses in the definition of a group with the more "minimal" statements:

3′. **Identity.** There exists a unique element, called the **unit** or **identity** and denoted by e, such that $ae = a$ for every element a in G.

4′. **Inverses.** Every element a in G has a unique inverse, denoted by a^{-1}, which is also in G, such that $a^{-1}a = e$.

[2]In abstract algebra (the theory of calculation), binary composition can be associative or non-associative. The most important non-associative algebras in physics are Lie algebras, which will be discussed later in this course.

The terms "multiplication," "product," and "unit" used in this definition are not meant to imply that the composition law corresponds to ordinary multiplication. The multiplication of two elements is only an abstract rule for combining an ordered pair of two group elements to obtain a third group element. The difference from ordinary multiplication becomes even more apparent from the fact that the composition law need not be commutative, i.e., the product ab need not equal ba for distinct group elements a and b. If a group does have a commutative composition law, it is said to be **commutative** or **Abelian**.

Despite the somewhat abstract tone of these comments, a moment's reflection leads to the realization that the structure of groups is ideally suited to the description of symmetry in physical systems. The group elements often correspond to coordinate transformations of either geometrical objects or of equations of motion, with the composition law corresponding to matrix multiplication or the usual composition law of functions,[3] so the associativity property is guaranteed.[4] If two operations each correspond to symmetry operations, then their product clearly must as well. The identity corresponds to performing no transformation at all and the inverse of each transformation corresponds to performing the transformation in reverse, which must exist for the transformation to be well-defined (cf. Example 2.4).

Example 2.1. Consider the set of integers,

$$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

with the composition rule being ordinary addition. The sum of any two integers is an integer, thus ensuring closure, addition is an associative operation, 0 is the identity, and the inverse of n is $-n$, which is clearly an integer. Thus, the integers form a group under addition. This group is denoted by Z (derived from the German word *Zahlen* for integers).

[3] For two functions $f(x)$ and $g(x)$, the application of f, followed by the application of g is $g[f(x)]$, and the application of g followed by the application of f is $f[g(x)]$.

[4] The associativity of linear operations in general, and matrices in particular, is discussed by Wigner in *Group Theory* (Academic, New York, 1959), along with other group properties.

Since the order in which two integers are added is immaterial, Z is an Abelian group. ▮

Example 2.2. The importance of the composition law for determining whether a set of elements forms a group can be seen by again considering the integers, but now with ordinary multiplication as the composition rule. The product of any two integers is again an integer, multiplication is associative, the unit is 1, but the inverse of n is $1/n$, which is *not* an integer if $n \neq 1$. Hence, the integers with ordinary multiplication do not form a group. ▮

Example 2.3. Consider the elements $\{1, -1\}$ under ordinary multiplication. This set is clearly closed under multiplication and associativity is manifestly satisfied. The unit element is 1 and each element is its own inverse. Hence, the set $\{1, -1\}$ is a two-element group under multiplication. ▮

Example 2.4. Consider the set of 2×2 matrices with real entries

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{2.1}$$

such that the determinant, $ad - bc$, is non-zero. The composition law is the usual rule for matrix multiplication:

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{pmatrix}.$$

To determine if this set of matrices forms a group, we must first show that the product of two matrices with non-zero determinant is also a matrix with non-zero determinant. This follows from that fact that for any pair of 2×2 matrices A and B, their determinants, denoted by $\det(A)$ and $\det(B)$, satisfy $\det(AB) = (\det A)(\det B)$. Associativity can be verified with a straightforward, but laborious, calculation. The identity is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and the inverse of (2.1) is

$$\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

which explains the requirement that $ad-bc \neq 0$. This group is denoted by GL(2,**R**), for *general linear* group of 2×2 matrices with real entries. Note that the elements of this group form a *continuous* set, so GL(2,**R**) is a continuous group. ∎

2.2 Permutation Groups

A permutation of n objects is a rearrangement of those objects. When combined with the usual rule for function composition for successive permutations (see below), these permutations are endowed with the structure of a group, which is denoted by S_n. At one time, permutation groups were the only groups studied by mathematicians and they maintain a special status in the subject through **Cayley's theorem**, which establishes a relationship between S_n and *every* group with n elements. In this section, we will examine the structure of S_3, both as an abstract group and as the symmetry group of an equilateral triangle.

The group S_3 is the set of all permutations of three distinguishable objects, where each element corresponds to a particular permutation of the three objects from a given reference order. Since the first object can be put into any one of three positions, the second object into either of two positions, and the last object only into the remaining position, there are $3\times2\times1=6$ elements in the set. These are listed below:

$$e=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix} \qquad a=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix} \qquad b=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$$

$$c=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix} \qquad d=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix} \qquad f=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$$

In this notation, the top line represents the initial, or reference, order of the objects and the bottom line represents the effect of the permutation. The composition law corresponds to performing successive

permutations and is carried out by rearranging the objects according to the first permutation and then using this as the reference order to rearrange the objects according to the second permutation. As an example, consider the product ad, where we will use the convention that operations are performed from right to left, i.e., permutation d is performed first, followed by permutation a. Element d permutes the reference order $(1, 2, 3)$ into $(3, 1, 2)$. Element a then permutes this by putting the first object in the second position, the second object into the first position, and leaves the third object in position three, i.e.,

$$a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix}.$$

Notice that it is only the permutation of the distinct objects, not their labelling, which is important for specifying the permutation. Hence,

$$ad = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = b\,,$$

An analogous procedure shows that

$$da = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = c\,,$$

which shows that the composition law is not commutative, so S_3 is a non-Abelian group.

A geometric realization of S_3 can be established by considering the symmetry transformations of an equilateral triangle (Fig. 2.1). The elements a, b, and c correspond to reflections through lines which intersect the vertices at 3, 1, and 2, respectively, and d and f correspond to clockwise rotations of this triangle by $\frac{2}{3}\pi$ and $\frac{4}{3}\pi$ radians, respectively. The effects of each of these transformations on the positions of the vertices of the triangle is identical with the corresponding element of S_3. Thus, there is a one-to-one correspondence between these transformations and the elements of S_3. Moreover, this correspondence is preserved by the composition laws in the two groups. Consider for example, the products ad and da calculated above for S_3. For the equilateral triangle, the product ad corresponds to a rotation followed by

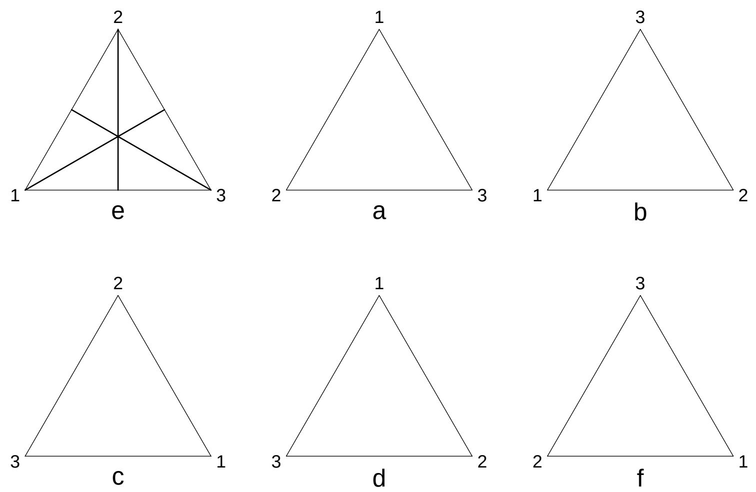
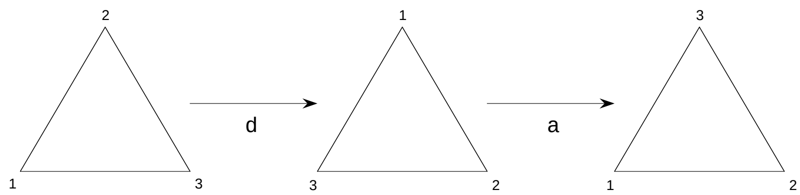


Figure 2.1: The symmetry transformations of an equilateral triangle labelled by the corresponding elements of S_3. The lines in the diagram corresponding to the identity are lines through which reflections of transformations a, b and c are taken. The transformations d and f are rotations.

a reflection. Thus, beginning with the standard order shown for the identity the successive application of these transformations is shown below:



By comparing with Fig. 2.1, we see that the result of these transformations is equivalent to the transformation b. Similarly, one can show that $da = c$ and, in fact, that all the products in S_3 are identical to those of the symmetry transformations of the equilateral triangle. Two such groups that have the same algebraic structure are said to be **isomorphic** to one another and are, to all intents and purposes, identical. This highlights the fact that it is the algebraic structure of the group

which is important, not any particular realization of the group. Further discussion of this point will be taken up in the next chapter.

2.3 Elementary Properties of Groups

The examples in the preceding section showed that all groups are endowed with several general properties. In this section, we deduce some additional properties which, although evident in particular examples, can be shown generally to follow from the properties of abstract groups.

Theorem 2.1. (Uniqueness of the identity) The identity element in a group G is unique.

Proof. Suppose there are two identity elements e and e' in G. Then, according to the definition of a group, we must have that

$$ae = a$$

and

$$e'a = a$$

for all a in G. Setting $a = e'$ in the first of these equations and $a = e$ in the second shows that

$$e' = e'e = e\,,$$

so $e = e'$. ∎

This theorem enables us to speak of *the* identity e of a group. The notation e is derived from the German word *Einheit* for unity.

Another property common to all groups is the cancellation of common factors within equations. This property owes its existence to the associativity of the group composition rule.

Theorem 2.2. (Cancellation) In a group G, the left and right cancellation laws hold, i.e., $ab = ac$ implies $b = c$ and $ba = ca$ implies $b = c$.

Proof. Suppose that $ab = ac$. Let a^{-1} be an inverse of a. Then, by left-multiplying by this inverse,

$$a^{-1}(ab) = a^{-1}(ac)$$

and invoking associativity,

$$(a^{-1}a)b = (a^{-1}a)c\,,$$

we obtain

$$eb = ec\,,$$

so $b = c$. Similarly, beginning with $ba = ca$ and right-multiplying by a^{-1} shows that $b = c$ in this case also. ∎

Notice that the proof of this theorem does not require the inverse of a group element to be unique; only the existence of *an* inverse was required. In fact, the cancellation theorem can be used to prove that inverses are, indeed, unique.

Theorem 2.3. (Uniqueness of inverses) For each element a in a group G, there is a unique element b in G such that $ab = ba = e$.

Proof. Suppose that there are two inverses b and c of a. Then $ab = e$ and $ac = e$. Thus, $ab = ac$, so by the Cancellation Theorem, $b = c$. ∎

As in the case of the identity of a group, we may now speak of *the* inverse of every element in a group, which we denote by a^{-1}. As was discussed in Sec. 2.1, this notation is borrowed from ordinary multiplication, as are most other notations for the group composition rule. For example, the n-fold product of a group element g with itself is denoted

by g^n. Similarly $g^n g^m = g^{n+m}$, which conforms to the usual rule of exponents for real numbers. However, there are some notable exceptions. For two group elements a and b, the equality of $(ab)^n$ and $a^n b^n$ does not generally hold. As the examples in Sec. 2.1 demonstrated, as long as this notation is interpreted in the context of the appropriate group composition rule, no confusion should arise.

2.4 Discrete and Continuous Groups

Groups are divided into two general categories: discrete and continuous. The basis definitions apply to both types of group, but the discussion of a number of properties depends sensitively on the discrete or continuous nature of the group. In this course, we will focus our attention on discrete groups first, to establish a conceptual base, and consider continuous groups later in the course.

2.4.1 Finite Groups

One of the most fundamental properties of a group G is number of elements contained in the group. This is termed the **order** of G and is denoted by $|G|$. The group Z of integers under addition, has infinite order and the order of S_3, the group of permutations of three objects, is 6. We will be concerned initially with *finite* groups which, apart from their applicability to a range of physical problems, have a number interesting arithmetic properties.

Finite groups also have properties which are not shared by either infinite or continuous groups. For example, if an element g of a finite group G is multiplied by itself enough times, the unit e is eventually recovered. Clearly, multiplying any element g by itself a number of times greater than $|G|$ *must* eventually lead to a recurrence of the product, since the number of distinct products is bounded from above by $|G|$. To show this explicitly, we denote a recurring product by a and write

$$a = g^p = g^q \,,$$

where $p = q + n$. Then, by using the associativity of the composition

law, $g^{q+n} = g^q g^n = g^n g^q$, so

$$g^p = g^q g^n = g^n g^q = g^q \,,$$

and, from the definition of the identity and its uniqueness, we conclude that

$$g^n = e \,.$$

Thus, the set of elements $g, g^2, g^3, \ldots$ represents a recurring sequence. The **order of an element** g, denoted by $|g|$, is the *smallest* value of k such that $g^k = e$. The **period** of such an element g is the collection of elements $\{e, g, g^2, \ldots, g^{k-1}\}$.

Example 2.5. Using S_3 as an example, $|a| = |b| = |c| = 2$ and $|d| = |f| = 3$. The corresponding periods are $\{e, a\}$, $\{e, b\}$, $\{e, c\}$, and $\{e, d, f = d^2\}$. ∎

Theorem 2.4. (Rearrangement Theorem) If $\{e, g_1, g_2, \ldots, g_n\}$ are the elements of a group G, and if g_k is an arbitrary group element, then the set of elements

$$Gg_k = \{eg_k, g_1g_k, g_2g_k, \ldots, g_ng_k\}$$

contains each group element once and only once.

Proof. The set Gg_k contains $|G|$ elements. Suppose two elements of Gg_k are equal: $g_ig_k = g_jg_k$. By the Cancellation Theorem, we must have that $g_i = g_j$. Hence, each group element appears once and only once in Gg_k, so the sets G and Gg_k are identical apart from a rearrangement of the order of the elements if g_k is not the identity. ∎

2.4.2 Multiplication Tables

One application of this theorem is in the representation of the composition law for a finite group as a **multiplication table**. Such a table is a square array with the rows and columns labelled by the elements of the group and the entries corresponding to the products, i.e., the element g_{ij} in the ith row and jth column is the product of the element g_i labelling that row and the element g_j labelling that column: $g_{ij} = g_i g_j$. To see how the construction of multiplication tables proceeds by utilizing only the abstract group properties, consider the simplest nontrivial group, that with two distinct elements $\{e, a\}$. We clearly must have the products $e^2 = e$ and $ea = ae = a$. The Rearrangement Theorem then requires that $a^2 = e$. The multiplication table for this group is shown below:

	e	a
e	e	a
a	a	e

Note that the entries of this table are symmetric about the main diagonal, which implies that this group is Abelian.

Now consider the group with three distinct elements: $\{e, a, b\}$. The only products which we must determine explicitly are ab, ba, a^2, and b^2 since all other products involve the unit e. The product ab cannot equal a or b, since that would imply that either $b = e$ or $a = e$, respectively. Thus, $ab = e$. The Cancellation Theorem then *requires* that $a^2 = e$, $b^2 = a$, and $ba = e$. The multiplication table for this group is shown below:

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Because the entries of this table are symmetric about the main diagonal, this group is also Abelian. Our procedure shows that *every* group with two or three elements *must* have the multiplication tables just

calculated, i.e., the algebraic structures of group with two and three elements are *unique*! Thus, we can speak of *the* group with two elements and *the* group with three elements. A similar procedure for groups with four elements $\{e, a, b, c\}$ yields *two* distinct multiplication tables (Problem Set 2). As a final example, the multiplication table for S_3 is shown below:

	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	e	d	f	b	c
b	b	f	e	d	c	a
c	c	d	f	e	a	b
d	d	c	a	b	f	e
f	f	b	c	a	e	d

As is immediately evident from this table, S_3 *not* Abelian.

2.5 Subgroups and Cosets

If, from a group G, we select a subset H of elements which themselves form a group under the same composition law, H is said to be a **subgroup** of G. According to this definition, the unit element $\{e\}$ forms a subgroup of G, and G is a subgroup of itself. These are termed *improper* subgroups. The determination of *proper* subgroups is one of the central concerns of group theory. In physical applications, subgroups arise in the description of symmetry-breaking, where a term is added to a Hamiltonian or a Lagrangian which lowers the symmetry to a subgroup of the original symmetry operations.

Example 2.6. The group S_3 has a number of proper subgroups: $\{e, a\}$, $\{e, b\}$, $\{e, c\}$, and $\{e, d, f\}$. The identification of these subgroups is most easily carried out by referring to the symmetry operations of an equilateral triangle (Fig. 2.1). ∎

If $H = \{e, h_1, h_2, \ldots, h_r\}$ is a subgroup of a group G, and g is an element of G, then the set

$$Hg = \{eg, h_1g, h_2g, \ldots, h_rg\}$$

is a **right coset** of H. Similarly, the set

$$gH = \{ge, gh_1, gh_2, \ldots, gh_r\}$$

is a **left coset** of H. A coset need not be a subgroup; it will be a subgroup only if g is an element of H.

Theorem 2.5. Two cosets of a subgroup either contain exactly the same elements or else have no common elements.

Proof. These cosets either have no common elements or have at least one common element. We will show that if there is a single in common, then all elements are common to both subgroups. Let Hg_1 and Hg_2 be two right cosets. If one common element of these cosets is $h_ig_1 = h_jg_2$, then

$$g_2g_1^{-1} = h_j^{-1}h_i$$

so $g_2g_1^{-1}$ is in H. But also contained in H are the elements

$$Hg_2g_1^{-1} = \{eg_2g_1^{-1}, h_1g_2g_1^{-1}, h_2g_2g_1^{-1}, \ldots, h_rg_2g_1^{-1}\}$$

since, according to the Rearrangement Theorem, each element of H appears once and only once in this sequence. Therefore, the elements of Hg_1 are identical to those of

$$(Hg_2g_1^{-1})g_1 = Hg_2(g_1^{-1}g_1) = Hg_2$$

so these two cosets have only common elements. ∎

Example 2.7. Consider again the group S_3 and its subgroup $H = \{e, a\}$ (Example 2.6). The right cosets of this subgroup are

$$\{e, a\}e = \{e, a\}, \quad \{e, a\}a = \{a, e\}, \quad \{e, a\}b = \{b, d\}$$

$$\{e, a\}c = \{c, f\}, \quad \{e, a\}d = \{d, b\}, \quad \{e, a\}f = \{f, c\}$$

We see that there are three distinct right cosets of $\{e, a\}$,

$$\{e, a\}, \qquad \{b, d\}, \qquad \{c, f\}$$

of which only the first is a subgroup (why?). Similarly, there are three *left* cosets of $\{e, a\}$:

$$\{e, a\}, \qquad \{c, d\}, \qquad \{b, f\}$$

Notice that the left and right cosets are *not* the same. ∎

Theorem 2.6 (Lagrange's theorem). The order of a subgroup H of a finite group G is a divisor of the order of G, i.e., $|H|$ divides $|G|$.

Proof. Cosets either have all elements in common or they are distinct (Theorem 2.5). This fact, combined with the Rearrangement Theorem, means that every element of the group must appear in exactly one distinct coset. Thus, since each coset clearly has the same number of elements, the number of distinct cosets, which is called the **index** of the subgroup, multiplied by the number of elements in the coset, is equal to the order of the group. Hence, since the order of the coset and the subgroup are equal, the order of the group divided by the order of the subgroup is equal to the number of distinct cosets, i.e., an integer. ∎

Example 2.8. The subgroup $\{e, a\}$ of S_3 is of order 2 and the subgroup $\{e, d, f\}$ is of order 3. Both 2 and 3 are divisors of $|S_3| = 6$. ∎

Lagrange's theorem identifies the allowable orders of the subgroups of a given group. But the converse of Lagrange's theorem is *not* generally valid, i.e., the orders of the subgroups of a group G need not span the divisors of G.

2.6 The Quotient Group

2.6.1 Conjugacy Classes

Two elements a and b of a group G are said to be **conjugate** if there is an element g in the group, called the conjugating element, such that $a = gbg^{-1}$. Conjugation is an example of what is called an **equivalence relation**, which is denoted by "$\equiv$," and is defined by three conditions:

1. $a \equiv a$ (reflexive).
2. If $a \equiv b$, then $b \equiv a$ (symmetric).
3. If $a \equiv b$ and $b \equiv c$, then $a \equiv c$ (transitive).

To show that conjugacy corresponds to an equivalence relation we consider each of these conditions in turn. By choosing $g = e$ as the conjugating element, we have that $a = eae^{-1} = a$, so $a \equiv a$. If $a \equiv b$, then $a = gbg^{-1}$, which we can rewrite as

$$g^{-1}ag = g^{-1}a(g^{-1})^{-1} = b$$

so $b \equiv a$, with g^{-1} as the conjugating element. Finally, to show transitivity, the relations $a \equiv b$ and $b \equiv c$ imply that there are elements g_1 and g_2 such that $b = g_1ag_1^{-1}$ and $c = g_2bg_2^{-1}$. Hence,

$$c = g_2bg_2^{-1} = g_2g_1ag_1^{-1}g_2^{-1} = (g_2g_1)a(g_2g_1)^{-1}$$

so c is conjugate to a with the conjugating element g_1g_2. Thus, conjugation fulfills the three conditions of an equivalency class.

One important consequence of equivalence is that it permits the assembly of **classes**, i.e., sets of equivalent quantities. In particular, a **conjugacy class** is the totality of elements which can be obtained from a given group element by conjugation. Group elements in the same conjugacy class have several common properties. For example, all elements of the same class have the same order. To see this, we begin with the definition of the order n of an element a as the smallest integer such that $a^n = e$. An arbitrary conjugate b of a is $b = gag^{-1}$. Hence,

$$b^n = \underbrace{(gag^{-1})(gag^{-1})\cdots(gag^{-1})}_{n\,\text{factors}} = ga^ng^{-1} = geg^{-1} = e$$

so b has the same order as a.

Example 2.9. The group S_3 has three classes: $\{e\}$, $\{a,b,c\}$, and $\{d,f\}$. As we discussed in Example 2.5, the order of a, b, and c is two, and the order of d and f is 3. The order of the unit element is 1 and is always in a class by itself. Notice that each class corresponds to a distinct kind of symmetry operation on an equilateral triangle. The operations a, b, and c correspond to reflections, while d and f correspond to rotations. In terms of operations in S_3, the elements d and f correspond to *cyclic* permutations of the reference order, e.g., $1 \to 2$, $2 \to 3$, and $3 \to 1$, while a, b, and c correspond to permutations which are not cyclic. ∎

2.6.2 Self-Conjugate Subgroups

A subgroup H of G is **self-conjugate** if the elements gHg^{-1} are identical with those of H for all elements g of G. The terms **invariant subgroup** and **normal subgroup** are also used. A group with no self-conjugate proper subgroups is called **simple**. If $gHg^{-1} = H$ for all g in G, then given an element h_1 in H, for any a, we can find an element h_2 in H such that $ah_1a^{-1} = h_2$, which implies that $ah_1 = h_2a$, or that $aH = Ha$. This last equality yields another definition of a self-conjugate subgroup as one whose left and right cosets are equal. From the definition of self-conjugacy and of classes, we can furthermore conclude that a subgroup H of G is self-conjugate if and only if it contains elements of G in complete classes, i.e., H contains either all or none of the elements of classes of G.

The cosets of a self-conjugate subgroup are themselves endowed with a group structure, with multiplication corresponding to an element-by-element composition of two cosets and discounting duplicate products. We show first that the multiplication of the elements of two right cosets of a conjugate subgroup yields another right coset. Let H be a self-conjugate subgroup of G and consider the two right cosets Ha and Hb. Then, the multiplication of Ha and Hb produces products of the form

$$h_iah_jb = h_i(ah_j)b$$

The product ah_j can be written as $h_k a$ for some h_k in H, since H is assumed to be self-conjugate. Thus, we have

$$h_i(ah_j)b = h_i(h_k a)b = (h_i h_k)(ab)$$

which is clearly an element of a right coset of H.

Example 2.10. Consider the subgroup $\{e, d, f\}$ of S_3. Right-multiplying this subgroup by each element of S_3 yields the right cosets of this subgroup:

$$\{e, d, f\}e = \{e, d, f\}, \quad \{e, d, f\}a = \{a, c, b\}, \quad \{e, d, f\}b = \{b, a, c\}$$

$$\{e, d, f\}c = \{c, b, a\}, \quad \{e, d, f\}d = \{d, f, e\}, \quad \{e, d, f\}f = \{f, e, d\}$$

Similarly, left-multiplying by each element of S_3 produces the left cosets of this subgroup:

$$e\{e, d, f\} = \{e, d, f\}, \quad a\{e, d, f\} = \{a, b, c\}, \quad b\{e, d, f\} = \{b, c, a\}$$

$$c\{e, d, f\} = \{c, a, b\}, \quad d\{e, d, f\} = \{d, f, e\}, \quad f\{e, d, f\} = \{f, e, d\}$$

Thus, since the right and left cosets of $\{e, d, f\}$ are the same, these elements form a self-conjugate subgroup of S_3 whose distinct cosets are $\{e, d, f\}$ and $\{a, b, c\}$. Multiplying these subgroups together and neglecting duplicate elements yields

$$\{e, d, f\}\{e, d, f\} = \{e, d, f\}, \quad \{e, d, f\}\{a, b, c\} = \{a, b, c\}$$

$$\{a, b, c\}\{e, d, f\} = \{a, b, c\}, \quad \{a, b, c\}\{a, b, c\} = \{e, d, f\}$$

∎

The **quotient group** (also called the **factor group**) of a self-conjugate subgroup is the collection of cosets, each being considered an element. The order of the factor group is equal to the index of the self-conjugate subgroup. With the notation used above, the quotient group is denoted by G/H.

Example 2.11. The cosets of the self-conjugate subgroup $\{e, d, f\}$ of S_3 are $\{e, d, f\}$ and $\{a, b, c\}$, so the order of the factor group is two. If we use the notation

$$\mathcal{E} = \{e, d, f\}, \qquad \mathcal{A} = \{a, b, c\} \tag{2.2}$$

for the elements of the factor group, we can use the results of Example 2.8 to construct the multiplication table for this group (shown below) from which see that $\mathcal{E}$ is the identity of the factor group, and $\mathcal{E}$ and

	$\mathcal{E}$	$\mathcal{A}$
$\mathcal{E}$	$\mathcal{E}$	$\mathcal{A}$
$\mathcal{A}$	$\mathcal{A}$	$\mathcal{E}$

$\mathcal{A}$ are their own inverses. Note that this multiplication table has the identical structure as the two-element group $\{e, a\}$ discussed in Sec. 2.4. ▮

2.7 Summary

In this chapter, we have covered only the most basic properties of groups. One of the remarkable aspects of this subject, already evident in some of the discussion here, is that the four properties that define a group, have such an enormous implication for the properties of groups, quite apart from their implications for physical applications, which will be explored throughout this course. A comprehensive discussion of the mathematical theory of groups, including many wider issues in both pure and applied mathematics, may be found in the book by Gallian.[5]

[5]J.A. Gallian, *Contemporary Abstract Algebra* 4th edn. (Houghton Mifflin, Boston, 1998).

Chapter 3

Representations of Groups

How can it be that mathematics, being after all a product of human thought which is independent of experience, is so admirably appropriate to the objects of reality?
—Albert Einstein

The structure of abstract groups developed in Chapter 2 forms the basis for the application of group theory to physical problems. Typically in such applications, the group elements correspond to symmetry operations which are carried out on spatial coordinates. When these operations are represented as linear transformations with respect to a coordinate system, the resulting matrices, together with the usual rule for matrix multiplication, form a group that is equivalent to the group of symmetry operations in a sense to be made precise later in this chapter. In essence, these matrices form what is called a *representation* of the symmetry group with each element corresponding to a particular matrix.

For applications to quantum mechanics, as we have seen in Section 1.2, the symmetry operations are performed on the Hamiltonian, whose invariance properties determine the symmetry group. The wavefunctions, which do not all share the symmetry of the Hamiltonian, will be seen to determine the representations of the symmetry group in the sense described above. These representations will, in turn, provide a classification scheme for the eigenfunctions of the Hamiltonian, in

analogous fashion to the identification of even and odd eigenfunctions in Section 1.2. The strength of the group-theoretic formalism that we will develop is that this procedure can be carried out in a systematic fashion for a Hamiltonian having any symmetry without undue computational effort.

In this chapter, we will set out the basic definitions that enable us to construct a mathematical definition of what we mean by a representation and discuss the basic types of representation. In the next chapter we will develop a number of remarkable properties of representations that lie at the heart of applications of discrete group theory to quantum mechanics.

3.1 Homomorphisms and Isomorphisms

Consider two finite groups G and G' with elements $\{e, a, b, \ldots\}$ and $\{e', a', b', \ldots\}$ and which need not be of the same order. Suppose there is a mapping ϕ between the elements of G and G' which preserves their composition rules, i.e., if $a' = \phi(a)$ and $b' = \phi(b)$, then

$$\phi(ab) = \phi(a)\phi(b) = a'b'$$

If the order of the two groups is the same, then this mapping is said to be an **isomorphism** and the two groups are **isomorphic** to one another. This is denoted by $G \approx G'$. If the order of the two groups is *not* the same, then the mapping is a **homomorphism** and the two groups are said to be **homomorphic**. Thus, an isomorphism is a one-to-one correspondence between two groups, while a homomorphism is a many-to-one correspondence. An isomorphism preserves the structure of the original group, but a homomorphism causes some of the structure of the original group to be lost. Both properties are reflected in the behavior of multiplication tables under these mappings. Homomorphisms and isomorphisms are not limited to finite groups nor even to groups with discrete elements.

Example 3.1. We saw in Sec. 2.2 that S_3 is isomorphic to the planar symmetry operations of an equilateral triangle, since there is a one-to-one correspondence between the elements of the two groups and they

have the same multiplication table. On the other hand, consider the correspondence between the elements of S_3 and the elements of the quotient group of S_3 discussed in Examples 2.9 and 2.11):

$$\{e,d,f\} \mapsto \{\mathcal{E}\}, \qquad \{a,b,c\} \mapsto \{\mathcal{A}\} \tag{3.1}$$

i.e. the mapping ϕ is defined by

$$\begin{array}{lll} \phi(e) = \mathcal{E} & \phi(d) = \mathcal{E} & \phi(f) = \mathcal{E} \\ \phi(a) = \mathcal{A} & \phi(b) = \mathcal{A} & \phi(c) = \mathcal{A} \end{array} \tag{3.2}$$

This is a *homomorphism* because three elements of S_3 correspond to a single element of the quotient group. To see that this mapping preserves multiplication, we rearrange the multiplication table of S_3 (Example 2.2) as follows:

$$\begin{array}{|c|ccc|ccc|} \hline & e & d & f & a & b & c \\ \hline e & e & d & f & a & b & c \\ d & d & f & e & c & a & b \\ f & f & e & d & b & c & a \\ \hline a & a & b & c & e & d & f \\ b & b & c & a & f & e & d \\ c & c & a & b & d & f & e \\ \hline \end{array} \quad \mapsto \quad \begin{array}{|c|c|c|} \hline & \mathcal{E} & \mathcal{A} \\ \hline \mathcal{E} & \mathcal{E} & \mathcal{A} \\ \hline \mathcal{A} & \mathcal{A} & \mathcal{E} \\ \hline \end{array}$$

where the mapping of the multiplication table onto the elements $\{\mathcal{E},\mathcal{A}\}$ is precisely that of a two-element group (cf. Example 2.9). This homomorphism clearly causes some of the structure of the original group to be lost. For example, S_3 is non-Abelian group, but the two-element group is Abelian.▮

3.2 Representations

A **representation** of dimension n of an abstract group G is a homomorphism or isomorphism between the elements of G and the group of nonsingular $n \times n$ matrices (i.e. $n \times n$ matrices with non-zero determinant) with complex entries and with ordinary matrix multiplication as the composition law (Example 2.4). An isomorphic representation

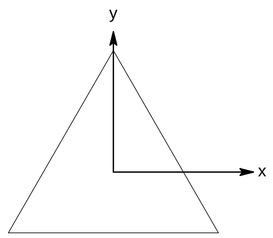


Figure 3.1: The coordinate system used to generate a two-dimensional representation of the symmetry group of the equilateral triangle. The origin of the coordinate system coincides with the geometric center of the triangle.

is called a **faithful** representation and a homomorphic representation is called an **unfaithful** representation.

According to this definition, if elements a and b of G are assigned matrices $D(a)$ and $D(b)$, then $D(a)D(b) = D(ab)$. The nonsingular nature of the matrices is required because inverses must be contained in the set (Example 2.4). Representations can also be comprised of numbers; the dimensionality of such representations is unity.

Example 3.2. Consider the following matrix representation of S_3 based on the correspondence with planar symmetry operations of an equilateral triangle:

$$\begin{aligned} e &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad a = \tfrac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad b = \tfrac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix} \\ c &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad d = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}, \quad f = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix} \end{aligned} \tag{3.3}$$

These matrices were generated by regarding each of the symmetry operations as a linear transformation in the coordinate system shown in Fig. 3.1. Matrices a, b, and c correspond to reflections, so their determinant is -1, while matrices d and f correspond to rotations, so their determinant is 1. These matrices form a faithful representation of S_3.

Consider now the following mapping between the elements of S_3 and the set $\{1,-1\}$:

$$\{e,f\} \mapsto \{1\}, \qquad \{a,b,c\} \mapsto \{-1\} \tag{3.4}$$

This is essentially a mapping between the elements of S_3 and the determinant of their matrix representation discussed above. Thus, the identity matrix e and the rotations d and f have determinants of 1, while the reflections a, b, and c have determinants of -1. The physical interpretation of this homomorphism is therefore as a mapping from the individual elements of S_3 to their parity, i.e., whether they change the orientation of the coordinate system (-1) or not (1).[1] Since the determinant provides less information about a transformation than its matrix representation, it is clear that some information about the group structure of S_3 is not preserved by this homomorphism.

Finally, we note that the mapping of *all* elements to unity,

$$\{e,a,b,c,d,f\} \mapsto 1$$

is a representation of any group, though clearly an unfaithful one. This is called the **identical representation**. In the present case, the identical representation corresponds to a mapping from the group element to the absolute value of the determinant. Since all of the transformations preserve the lengths of vectors, any product of these transformations does so as well. ∎

Representations of groups are important in quantum mechanics for several reasons. First, the eigenfunctions of a Hamiltonian will transform under the symmetry operations of that Hamiltonian according to a particular representation of that group. Second, quantum mechanical operators are often written in terms of their matrix elements, so it is convenient to write symmetry operations in the same kind of matrix representation. Moreover, the evaluation of these matrix elements may sometimes be simplified by identifying the appropriate selection rules

[1] In terms of the operations of S_3, even parity corresponds to an even number of pairwise interchanges, while odd parity corresponds to an odd number of such interchanges.

(Section 1.2). Finally, the algebra of matrices is generally simpler to carry out than abstract symmetry operations. Thus, in the next section, we discuss some of the important properties of matrix representations of groups.

3.3 Reducible and Irreducible Representations

The definition of a representation provides for considerable flexibility in constructing matrix representations, which is manifested in several ways, but also indicates that representations are not unique. We consider some examples.

Given a matrix representation

$$\Big\{D(e), D(a), D(b), \ldots\Big\}$$

of an abstract group with elements $\{e, a, b, \ldots\}$, we can obtain a new set of matrices which also form a representation by performing a transformation known variously as a **similarity**, **equivalence**, or **canonical** transformation (cf. Sec. 2.6):

$$\Big\{BD(e)B^{-1}, BD(a)B^{-1}, BD(b)B^{-1}, \ldots\Big\} \tag{3.5}$$

Such transformations arise quite naturally, for example, in carrying out a change of basis for a set of matrices. Thus, suppose one begins with the matrix equation $\boldsymbol{b} = A\boldsymbol{a}$ relating two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ through a transformation A. If we now wish to express this equation in another basis which is obtained from the original basis by applying a transformation B, we can write

$$B\boldsymbol{b} = BA\boldsymbol{a} = BAB^{-1}B\boldsymbol{a}$$

so in the new basis, our original equation becomes

$$\boldsymbol{b}' = A'\boldsymbol{a}'$$

where $\boldsymbol{b}' = B\boldsymbol{b}$, $\boldsymbol{a}' = B\boldsymbol{a}$, and $A' = BAB^{-1}$. A similarity transformation can therefore be interpreted as a sequence of transformations

involving first a transformation to the original basis (B^{-1}), then performing the transformation A, and finally transforming back to the new basis (B). Referring back to the discussion in Section 2.6 on conjugacy classes, we see that group elements in the same conjugacy class represent the same type of transformation (e.g., reflection or rotation) which can be transformed into one another by particular symmetry operations.

Suppose we have representations of dimensions m and n. We can construct a representation of dimension $m+n$ by forming block-diagonal matrices:

$$\left\{\begin{bmatrix} D(e) & 0 \\ 0 & D'(e) \end{bmatrix}, \begin{bmatrix} D(a) & 0 \\ 0 & D'(a) \end{bmatrix}, \begin{bmatrix} D(b) & 0 \\ 0 & D'(b) \end{bmatrix}, \ldots\right\} \tag{3.6}$$

where $\{D(e), D(a), D(b), \ldots\}$ is an n-dimensional representation and $\{D'(e), D'(a), D'(b), \ldots\}$ an m-dimensional representation of the group G, and the symbol 0 is an $n\times m$ or an $m\times n$ zero matrix, as required by its position in the supermatrix. Each of the $m+n$-dimensional matrices formed in this manner is called a **direct sum** of the n- and m-dimensional component matrices. The direct sum is denoted by "$\oplus$" to distinguish it from the ordinary addition of two matrices. Thus, we can write the representation in (3.6) as

$$\left\{D(e) \oplus D'(e), D(a) \oplus D'(a), D(b) \oplus D'(b), \cdots\right\}$$

The two representations that form this direct sum can be either distinct or identical and, of course, the block-diagonal form can be continued indefinitely simply by incorporating additional representations in diagonal blocks. However, in all such constructions, we are not actually generating anything intrinsically new; we are simply reproducing the properties of known representations. Thus, although representations are a convenient way of associating matrices with group elements, the freedom we have in constructing representations, exemplified in (3.5) and (3.6), does not readily demonstrate that these matrices embody any intrinsic characteristics of the group they represent. Accordingly, we now describe a way of classifying equivalent representations and then introduce a refinement of our definition of a representation.

To overcome the problem of nonuniqueness posed by representations that are related by similarity transformations we consider the sum of the diagonal elements of an $n \times n$ matrix A, called the **trace** of A and by "tr":

$$\mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii}$$

The utility of the trace stems from its *invariance* under similarity transformations, i.e.,

$$\mathrm{tr}(A) = \mathrm{tr}(BAB^{-1})$$

The importance of this invariance, the proof of which is discussed in Problem Set 4, is that, although there is an infinite variety of representations related by similarity transformations, each such representation has the same set of traces associated with each of its elements.

But working with the trace alone does not alleviate the nonuniqueness of representations posed by (3.6). To address this issue, we introduce the concept of an *irreducible* representation. Representations such as those in (3.6) are termed *reducible* because they are the direct sum of two (or more) representations. We could, of course, perform a similarity transformation to obtain a representation that is not in block form, but the representation so obtained is still deemed to be reducible because it was obtained from matrices which originally were in block form. Based on these considerations, we define reducible and irreducible representations as follows:

Definition. If the *same* similarity transformation brings all of the matrices of a representation into the same block form (by which we mean matrices of the same dimension in the same positions), then this representation is said to be **reducible**. Otherwise, the representation is said to be **irreducible**.

Thus, irreducible representations cannot be expressed in terms of representations of lower dimensionality. One-dimensional representations are, by definition, always irreducible. Determining the irreducible

representations of groups is one of the central issues to be covered in the following chapters.

Example 3.4. All of the representations of S_3 discussed in Example 3.2 are *irreducible*. This is clear for the identical representation and for the representation in (3.4), since they are composed of numbers. But we can use these representations to construct the following manifestly *reducible* representation of S_3:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad d = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The representation in (3.3) is *irreducible*. There is no similarity transformation that will bring all of the matrices into block-diagonal form which, for the case here, means simple diagonalization. The easiest way to see this is from the point of view of the commutability of two matrices. If two matrices can be brought into diagonal form by the same similarity transformation, then they commute. As diagonal matrices, they certainly commute, so they must also commute in their original form. But a glance at the multiplication table for these matrices (recall that they are a *faithful* representation of S_3) in Example 2.2 shows that they do not all commute. Hence, they cannot all be simultaneously diagonalized, so this representation is *irreducible*. ∎

3.4 Unitary Representations

Representations of groups are useful because of orthogonality theorems which we will prove in the next chapter. As background to that discussion, we will prove in this section an important result about the unitarity of representations. But we first review some general properties of matrices.

We begin by considering the transformation of an $n \times n$ matrix A with entries A_{ij}, $i, j = 1, 2, \ldots, n$, under the action of various opera-

tions. The *complex conjugate* of A, denoted by A^*, has entries which are the complex conjugates of the corresponding entries of A:

$$(A^*)_{ij} = (A_{ij})^* \tag{3.7}$$

The *transpose* of A, denoted by $A^{\rm t}$, has its rows and columns interchanged with respect to those of A:

$$(A^{\rm t})_{ij} = A_{ji} \tag{3.8}$$

When applied to vectors, the transpose transforms a row vector into a column vector and *vice versa*. The transpose of a product of matrices $A, B, C, \ldots$ is

$$(ABC\cdots)^{\rm t} = \cdots C^{\rm t} B^{\rm t} A^{\rm t} \tag{3.9}$$

i.e., the order of matrix multiplication is *reversed*. This can be proven easily from the definition (3.8). Finally, the *adjoint* or **Hermitian conjugate** of A, denoted by $A^\dagger$, is the transposed conjugate of A, i.e.

$$(A^\dagger)_{ij} = (A_{ji})^* \tag{3.10}$$

In common with the transpose, the application of the Hermitian conjugate to a product of matrices $A, B, C, \ldots$ can be expressed as a product of Hermitian conjugates of the individual matrices, but with the order reversed:

$$(ABC\cdots)^\dagger = \cdots C^\dagger B^\dagger A^\dagger \tag{3.11}$$

3.4.1 Hermitian and Orthogonal Matrices

A matrix A is **Hermitian** if

$$A^\dagger = A \tag{3.12}$$

Hermitian matrices and Hermitian operators are familiar from quantum mechanics, where their properties of having real eigenvalues and orthogonal eigenvectors are of fundamental importance. A matrix A is **orthogonal** if its transpose is its inverse:

$$A^{\rm t} A = A A^{\rm t} = I \tag{3.13}$$

where I is the $n \times n$ unit matrix. In terms of matrix components, this condition reads

$$\sum_{k=1}^{n} a_{ki} a_{kj} = \sum_{k=1}^{n} a_{ik} a_{jk} = \delta_{ij} \tag{3.14}$$

where

$$\delta_{ij} = \begin{cases} 0, & \text{if } i \neq j; \\ 1, & \text{if } i = j \end{cases} \tag{3.15}$$

is the **Kronecker delta**. Thus, the rows of an orthogonal matrix are mutually orthogonal, as are the columns. The consequences of the orthogonality of a transformation matrix can be seen by examining the effect of applying an orthogonal matrix A to two n-dimensional vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, yielding vectors $\boldsymbol{u}'$ and $\boldsymbol{v}'$:

$$\boldsymbol{u}' = A\boldsymbol{u}, \qquad \boldsymbol{v}' = A\boldsymbol{v}$$

We now take the scalar, or 'dot,' product between $\boldsymbol{u}'$ and $\boldsymbol{v}'$:

$$(\boldsymbol{u}', \boldsymbol{v}') = (\boldsymbol{u}')^{\mathrm{t}}\boldsymbol{v}' = (A\boldsymbol{u})^{\mathrm{t}} A\boldsymbol{v} = \boldsymbol{u}^{\mathrm{t}} A^{\mathrm{t}} A\boldsymbol{v} = \boldsymbol{u}^{\mathrm{t}}\boldsymbol{v} = (\boldsymbol{u}, \boldsymbol{v}) \tag{3.16}$$

where we have used (3.11) and the fact that A is orthogonal. This shows that the relative orientations and the lengths of vectors are preserved by orthogonal transformations. Such transformations are either rigid rotations, which preserve the "handedness" (i.e., left or right) of a coordinate system, and are called **proper** rotations, or reflections, which reverse the "handedness" of a coordinate system, and are called **improper** "rotations."

3.4.2 Unitary Matrices

A third type of matrix, called **unitary**, has the property that

$$A^{\dagger} A = A A^{\dagger} = I \tag{3.17}$$

By writing this condition in terms of matrix components,

$$\sum_{k=1}^{n} a_{ki}^{*} a_{kj} = \sum_{k=1}^{n} a_{ik} a_{jk}^{*} = \delta_{ij} \tag{3.18}$$

we see that, in common with orthogonal matrices, the rows and columns of a unitary matrix are orthogonal, but with respect to a different scalar product. For two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, this scalar product is defined as

$$(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^\dagger \boldsymbol{v} \tag{3.19}$$

This generalizes the familiar dot product to complex vectors in n dimensions. We can now show, by proceeding as above, that unitary transformations leave the scalar product invariant:

$$(\boldsymbol{u}', \boldsymbol{v}') = (\boldsymbol{u}')^\dagger \boldsymbol{v}' = (A\boldsymbol{u})^\dagger A\boldsymbol{v} = \boldsymbol{u}^\dagger A^\dagger A\boldsymbol{v} = \boldsymbol{u}^\dagger \boldsymbol{v} = (\boldsymbol{u}, \boldsymbol{v}) \tag{3.20}$$

The property of unitarity, when applied to operators, is of immense importance in quantum mechanics because it enables changes of bases to be performed while preserving the orthogonality of bases and, thus, the overlap between wavefunctions. In this sense, unitary matrices are associated with proper and improper "rotations," in analogy with orthogonal matrices.

3.4.3 Diagonalization of Hermitian Matrices*

Let H be an $n \times n$ Hermitian matrix. The eigenvalue equation for this matrix is

$$H\boldsymbol{a} = \lambda \boldsymbol{a} \tag{3.21}$$

By writing this equation as

$$(H - \lambda I)\boldsymbol{a} = 0 \tag{3.22}$$

the eigenvalue equation in (3.21) has nontrivial solutions for $\boldsymbol{a}$ if and only if the determinant of the matrix of coefficients in (3.22) vanishes:

$$\det(H - \lambda I) = 0 \tag{3.23}$$

The expansion of the determinant leads to an nth-order polynomial in λ whose solution yields the n (not necessarily distinct) eigenvalues of H: $\lambda_1, \lambda_2, \ldots, \lambda_n$.

We now show that the eigenvectors of H which correspond to *distinct* eigenvalues are orthogonal. Consider the eigenvalue equations for

two eigenvectors $\boldsymbol{a}$ and $\boldsymbol{b}$ corresponding to distinct eigenvalues λ and μ, respectively:

$$H\boldsymbol{a} = \lambda \boldsymbol{a} \tag{3.24}$$

$$H\boldsymbol{b} = \mu \boldsymbol{b} \tag{3.25}$$

We now take the scalar product between $\boldsymbol{b}$ and (3.24) and that between (3.25) and $\boldsymbol{a}$:

$$\boldsymbol{b}^\dagger (H\boldsymbol{a}) = \lambda \boldsymbol{b}^\dagger \boldsymbol{a} \tag{3.26}$$

$$(\boldsymbol{b}^\dagger H^\dagger)\boldsymbol{a} = \mu \boldsymbol{b}^\dagger \boldsymbol{a} \tag{3.27}$$

Subtracting (3.27) from (3.26), and using the fact that H is Hermitian yields

$$(\lambda - \mu)\boldsymbol{b}^\dagger \boldsymbol{a} = \boldsymbol{b}^\dagger H \boldsymbol{a} - \boldsymbol{b}^\dagger H^\dagger \boldsymbol{a} = 0 \tag{3.28}$$

which, since $\lambda \neq \mu$, implies that $\boldsymbol{b}^\dagger \boldsymbol{a} = 0$, i.e., that $\boldsymbol{a}$ and $\boldsymbol{b}$ are orthogonal. If λ and μ are not distinct, we must use a Gram–Schmidt procedure to explicitly construct an orthogonal set of eigenvectors associated with the degenerate eigenvalue. Thus, the eigenvectors of a Hermitian matrix can always be chosen to form an orthogonal set.

Consider the matrix U whose columns are the eigenvectors of H:

$$U = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$$

We can then write (3.21) in a form that subsumes *all* the eigenvectors of H as follows:

$$HU = UD \tag{3.29}$$

where D is the diagonal matrix whose entries are the eigenvalues of H:

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{pmatrix}$$

Since the rows of U are composed of the (orthogonal) eigenvectors of H, it has the property that (cf. 3.18)

$$U^\dagger U = UU^\dagger = I$$

i.e., $U^{-1} = U^\dagger$, so U is *unitary.* Hence, we can rewrite (3.29) as

$$U^{-1}HU = U^\dagger HU = D$$

We have proven the following theorem:

Theorem 3.1. Any Hermitian matrix can be diagonalized by an appropriate unitary transformation.

This theorem will be used in the next section to prove an important result concerning the existence of unitary group representations.

3.4.4 Transformation to Unitary Representations

We have seen in Section 3.3 that there is considerable flexibility in constructing group representations. In this section, we take a first step in restricting this freedom by showing that any representation can be expressed entirely in terms of unitary matrices. Quite apart from the convenient properties of unitary matrices discussed in Section 3.4.2, this theorem allows to think of group representations as proper and improper complex "rotations."

Theorem 3.2. Every representation can be brought into unitary form by a similarity transformation.

Proof. Let $\{A_1, A_2, \ldots, A_{|G|}\}$ be a d-dimensional representation of a group G, i.e., the A_α are a set of $|G|$ $d \times d$ matrices with nonvanishing determinants. From these matrices we form a matrix H given by the sum

$$H = \sum_{\alpha=1}^{|G|} A_\alpha A_\alpha^\dagger$$

This matrix is Hermitian because, using the property (3.11),

$$H^\dagger = \sum_\alpha (A_\alpha A_\alpha^\dagger)^\dagger = \sum_\alpha A_\alpha A_\alpha^\dagger = H$$

According to Theorem 3.1, any Hermitian matrix can be diagonalized by some unitary transformation U. Denoting the diagonalized form of H by D, we have $D = U^\dagger H U$, which enables us to write D as

$$D = \sum_\alpha U^\dagger A_\alpha A_\alpha^\dagger U = \sum_\alpha (U^\dagger A_\alpha U)(U^\dagger A_\alpha^\dagger U) = \sum_\alpha (U^\dagger A_\alpha U)(U^\dagger A_\alpha U)^\dagger$$

By introducing the notation $\tilde{A}_\alpha = U^\dagger A_\alpha U$, we can write the last equation in a more concise form as

$$D = \sum_\alpha \tilde{A}_\alpha \tilde{A}_\alpha^\dagger \tag{3.30}$$

The diagonal elements of D are real, because

$$\begin{aligned} D_{kk} &= \sum_\alpha \sum_j (\tilde{A}_\alpha)_{kj} (\tilde{A}^\dagger{}_\alpha)_{jk} \\ &= \sum_\alpha \sum_j (\tilde{A}_\alpha)_{kj} (\tilde{A}_\alpha)^*_{kj} \\ &= \sum_\alpha \sum_j \left| (\tilde{A}_\alpha)_{kj} \right|^2 \end{aligned}$$

for $k = 1, 2, \ldots, d$, and positive, because the summation over j includes a diagonal element of the identity, which is a $d \times d$ unit matrix, and hence is equal to unity. Thus, the diagonal matrix $D^{1/2}$,

$$D^{1/2} = \begin{pmatrix} D_{11}^{1/2} & 0 & \cdots & 0 \\ 0 & D_{22}^{1/2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_{dd}^{1/2} \end{pmatrix}$$

and $D^{-1/2}$, which is given by an analogous expression, both have positive entries.

We now form the matrices

$$B_\alpha = D^{-1/2}\tilde{A}_\alpha D^{1/2}$$

from which we obtain the corresponding Hermitian conjugates:

$$B_\alpha^\dagger = (D^{-1/2}\tilde{A}_\alpha D^{1/2})^\dagger = D^{1/2}\tilde{A}^\dagger{}_\alpha D^{-1/2}$$

We will now demonstrate that the B_α are unitary by first showing that the product $B_\alpha B_\alpha^\dagger$ is equal to the identity matrix. The product $B_\alpha B_\alpha^\dagger$ is given by

$$\begin{aligned} B_\alpha B_\alpha^\dagger &= \Big(D^{-1/2}\tilde{A}_\alpha D^{1/2}\Big)\Big(D^{1/2}\tilde{A}^\dagger{}_\alpha D^{-1/2}\Big) \\ &= D^{-1/2}\tilde{A}_\alpha D\tilde{A}^\dagger{}_\alpha D^{-1/2} \end{aligned}$$

The definition of D in (3.30) and the associativity of matrix multiplication allow us to write this expression as

$$\begin{aligned} B_\alpha B_\alpha^\dagger &= D^{-1/2}\sum_j \tilde{A}_\alpha \tilde{A}_\beta \tilde{A}^\dagger{}_\beta \tilde{A}^\dagger{}_\alpha D^{-1/2} \\ &= D^{-1/2}\sum_j (\tilde{A}_\alpha \tilde{A}_\beta)(\tilde{A}_\alpha \tilde{A}_\beta)^\dagger D^{-1/2} \end{aligned}$$

Since the A_α are a representation of G, then so are the $\tilde{A}_\alpha$ (Problem 3, Problem Set 4). Hence, the product $\tilde{A}_\alpha \tilde{A}_\beta$ is another matrix $\tilde{A}_\gamma$ in this representation. Moreover, according to the Rearrangement Theorem, the sum over all β means that the set of $\tilde{A}_\gamma$ obtained from these products contains the matrix corresponding to each group element once and only once. Thus,

$$B_\alpha B_\alpha^\dagger = D^{-1/2}\underbrace{\sum_\gamma \tilde{A}_\gamma \tilde{A}_\gamma^\dagger}_{D} D^{-1/2} = I$$

where I is the $d \times d$ unit matrix. This result can also be used to show that $B_\alpha^\dagger B_\alpha = I$. Thus, the B_α, which are obtained from the original representation by a similarity transformation,

$$B_\alpha = D^{-1/2}U^{-1}A_\alpha U D^{1/2} = (UD^{1/2})^{-1}A_\alpha(UD^{1/2})$$

form a *unitary* representation of G. Hence, without any loss of generality, we may always assume that a representation is unitary. ∎

3.5 Summary

The main concepts introduced in this chapter are faithful and unfaithful representations, based on isomorphic and homomorphic mappings, respectively, reducible and irreducible representations, and the fact that we may confine ourselves to unitary representations of groups. In the next chapter we will focus on irreducible representations, both faithful and unfaithful, since these cannot be decomposed into representations of lower dimension and are, therefore, "intrinsic" to a symmetry group, since *all* reducible representations will be shown to be composed of direct sums of irreducible representations. Irreducible representations occupy a special place in group theory because they can be classified for a given symmetry group solely according to their traces and dimension.

Chapter 4

Properties of Irreducible Representations

Algebra is generous; she often gives more than is asked of her.
—Jean d'Alembert

We have seen in the preceding chapter that a reducible representation can, through a similarity transformation, be brought into block-diagonal form wherein each block is an irreducible representation. Thus, irreducible representations are the basic components from which all representations can be constructed. But the identification of whether a representation is reducible or irreducible is a time-consuming task if it relies solely on methods of linear algebra.[1] In this chapter, we lay the foundation for a more systematic approach to this question by deriving the fundamental theorem of representation theory, called the Great Orthogonality Theorem. The utility of this theorem, and its central role in the applications of group theory to physical problems, stem from the fact that it leads to simple criteria for determining irreducibility and provides a direct way of identifying the number of inequivalent representations for a given group. This theorem is based on two lemmas of Schur, which are the subjects of the first two sections of this chapter.

[1] K. Hoffman and R. Kunze, *Linear Algebra* 2nd edn (Prentice–Hall, Englewood Cliffs, New Jersey, 1971), Ch. 6,7.

4.1 Schur's First Lemma

Schur's two lemmas are concerned with the properties of matrices that commute with all of the matrices of a irreducible representations. The first lemma addresses the properties of matrices which commute with a given *irreducible* representation:

Theorem 4.1 (Schur's First Lemma). A non-zero matrix which commutes with all of the matrices of an irreducible representation is a constant multiple of the unit matrix.

Proof. Let $\{A_1, A_2, \ldots, A_{|G|}\}$ be the matrices of a d-dimensional irreducible representation of a group G, i.e., the A_α are $d \times d$ matrices which cannot all be brought into block-diagonal form by the same similarity transformation. According to Theorem 3.2, we can take these matrices to be unitary without any loss of generality. Suppose there is a matrix M that commutes with all of the A_α:

$$MA_\alpha = A_\alpha M \tag{4.1}$$

for $\alpha = 1, 2, \ldots, |G|$. By taking the adjoint of each of these equations, we obtain

$$A_\alpha^\dagger M^\dagger = M^\dagger A_\alpha^\dagger \,. \tag{4.2}$$

Since the A_α are unitary, $A_\alpha^\dagger = A_\alpha^{-1}$, so multiplying (4.2) from the left and right by A_α yields

$$M^\dagger A_\alpha = A_\alpha M^\dagger \,, \tag{4.3}$$

which demonstrates that, if M commutes with every matrix of a representation, then so does $M^\dagger$. Therefore, given the commutation relations in (4.1) and (4.3) any linear combination of M and $M^\dagger$ also commute with these matrices:

$$(aM + bM^\dagger)A_\alpha = A_\alpha(aM + bM^\dagger)\,,$$

where a and b are any complex constants. In particular, the linear combinations

$$H_1 = M + M^\dagger, \qquad H_2 = i(M - M^\dagger)$$

yield *Hermitian* matrices: $H_i = H_i^\dagger$ for $i = 1, 2$. We will now show that a Hermitian matrix which commutes with all the matrices of an irreducible representation is a constant multiple of the unit matrix. It then follows that M is also such a matrix, since

$$M = \tfrac{1}{2}(H_1 - iH_2) \tag{4.4}$$

The commutation between a general Hermitian matrix H and the A_α is expressed as

$$HA_\alpha = A_\alpha H\,. \tag{4.5}$$

Since H is Hermitian, there is a unitary matrix U which transforms H into a diagonal matrix D (Theorem 3.1):

$$D = U^{-1}HU\,.$$

We now perform the same similarity transformation on (4.5):

$$\begin{aligned} U^{-1}HA_iU &= U^{-1}HUU^{-1}A_iU \\ &= U^{-1}A_iHU = U^{-1}A_iUU^{-1}HU \end{aligned}$$

By defining $\tilde{A}_\alpha = U^{-1}A_\alpha U$, the transformed commutation relation (4.5) reads

$$D\tilde{A}_\alpha = \tilde{A}_\alpha D\,. \tag{4.6}$$

Using the fact that D is a diagonal matrix, i.e., that its matrix elements are $D_{ij} = D_{ii}\delta_{ij}$, where δ_{ij} is the Kronecker delta, the (m, n)th matrix element of the left-hand side of this equation is

$$(D\tilde{A}_\alpha)_{mn} = \sum_k D_{mk}(\tilde{A}_\alpha)_{kn} = \sum_k D_{mm}\delta_{mk}(\tilde{A}_\alpha)_{kn} = D_{mm}(A_\alpha)_{mn}\,.$$

Similarly, the corresponding matrix element on the right-hand side is

$$(\tilde{A}_\alpha D)_{mn} = \sum_k (\tilde{A}_\alpha)_{mk}D_{kn} = \sum_k (\tilde{A}_\alpha)_{mk}D_{nn}\delta_{kn} = (\tilde{A}_\alpha)_{mn}D_{nn}\,.$$

Thus, after a simple rearrangement, the (m, n)th matrix element of (4.6) is

$$(\tilde{A}_\alpha)_{mn}(D_{mm} - D_{nn}) = 0\,. \tag{4.7}$$

There are three cases that we must consider to understand the implications of this equation.

Case I. Suppose that all of the diagonal elements of D are distinct: $D_{mm} \neq D_{nn}$ if $m \neq n$. Then, (4.7) implies that

$$(\tilde{A}_\alpha)_{mn} = 0, \qquad m \neq n\,,$$

i.e., the off-diagonal elements of $\tilde{A}_\alpha$ must vanish, these are *diagonal* matrices and, therefore, according to the discussion in Section 3.3, they form a *reducible* representation composed of d one-dimensional representations. Since the $\tilde{A}_i$ are obtained from the A_i by a similarity transformation, the A_i themselves form a *reducible* representation.

Case II. If all of the diagonal elements of D are equal, i.e. $D_{mm} = D_{nn}$ for *all* m and n, then D is proportional to the unit matrix. The $(\tilde{A}_\alpha)_{mn}$ are not required to vanish for *any* m and n. Thus, only this case is consistent with the requirement that the A_α form an irreducible representation. If D is proportional to the unit matrix, then so is $H = UDU^{-1}$ and, according to (4.4), the matrix M is as well.

Case III. Suppose that the first p diagonal entries of D are equal, but the remaining entries are distinct from these and from each other: $D_{11} = D_{22} = \cdots = D_{pp}$, $D_{mm} \neq D_{nn}$ otherwise. The $(\tilde{A}_\alpha)_{mn}$ must vanish for any pair of unequal diagonal entries. These correspond to the cases where *both* m and n lie in the range $1, 2, \ldots, p$ and where m and n are equal and both greater than p, so *all* the $\tilde{A}_i$ all have the following general form:

$$\tilde{A}_i = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix},$$

where B_1 is a $p \times p$ matrix and B_2 is a $(p-d)\times(p-d)$ *diagonal* matrix. Thus, the $\tilde{A}_i$ are block diagonal matrices and are, therefore, *reducible*.

We have shown that if a matrix that not a multiple of the unit matrix commutes with all of the matrices of a representation, then that representation is necessarily *reducible* (Cases I and III). Thus, if a non-zero matrix commutes with all of the matrices of an *irreducible* representation (Case III), that matrix must be a multiple of the unit matrix. This proves Schur's lemma. ∎

4.2 Schur's Second Lemma

Schur's first lemma is concerned with the commutation of a matrix with a given irreducible representation. The second lemma generalizes this to the case of commutation with two distinct irreducible representations which may have different dimensionalities. Its statement is as follows:

Theorem 4.2 (Schur's Second Lemma). Let $\{A_1, A_2, \ldots, A_{|G|}\}$ and $\{A'_1, A'_2, \ldots, A'_{|G|}\}$ be two irreducible representations of a group G of dimensionalities d and d', respectively. If there is a matrix M such that

$$MA_\alpha = A'_\alpha M$$

for $\alpha = 1, 2, \ldots, |G|$, then if $d = d'$, either $M = 0$ or the two representations differ by a similarity transformation. If $d \neq d'$, then $M = 0$.

Proof. Given the commutation relation between M and the two irreducible representations,

$$MA_\alpha = A'_\alpha M\,, \tag{4.8}$$

we begin by taking the adjoint:

$$A^\dagger_\alpha M^\dagger = M^\dagger A'^\dagger_\alpha\,. \tag{4.9}$$

Since, according to Theorem 3.2, the A_α may be assumed to be unitary, $A^\dagger_\alpha = A^{-1}_\alpha$, so (4.9) becomes

$$A^{-1}_\alpha M^\dagger = M^\dagger A'^{-1}_\alpha\,. \tag{4.10}$$

By multiplying this equation from the left by M,

$$MA^{-1}_\alpha M^\dagger = MM^\dagger A'^{-1}_\alpha\,,$$

and utilizing the commutation relation (4.8) to write

$$MA^{-1}_\alpha = A'^{-1}_\alpha M\,,$$

we obtain

$$A'^{-1}_\alpha MM^\dagger = MM^\dagger A'^{-1}_\alpha\,.$$

Thus, the $d' \times d'$ matrix $MM^\dagger$ commutes with all the matrices of an irreducible representation. According to Schur's First Lemma, $MM^\dagger$ must therefore be a constant multiple of the unit matrix,

$$MM^\dagger = cI\,, \tag{4.11}$$

where c is a constant. We now consider individual cases.

Case I. $d = d'$. If $c \neq 0$, Eq. (4.11) implies that[2]

$$M^{-1} = \frac{1}{c}M^\dagger\,.$$

Thus, we can rearrange (4.8) as

$$A_\alpha = M^{-1}A'_\alpha M\,,$$

so our two representations are related by a similarity transformation and are, therefore, equivalent.

If $c = 0$, then $MM^\dagger = 0$. The (i,j)th matrix element of this product is

$$(MM^\dagger)_{ij} = \sum_k M_{ik}(M^\dagger)_{kj} = \sum_k M_{ik}M^*_{jk} = 0\,.$$

By setting $i = j$, we obtain

$$\sum_k M_{ik}M^*_{ik} = \sum_k |M_{ik}|^2 = 0\,,$$

which implies that $M_{ik} = 0$ for *all* i and k, i.e., that M is the zero matrix. This completes the first part of the proof.

Case II. $d \neq d'$. We take $d < d'$. Then M is a *rectangular* matrix with d columns and d' rows:

$$M = \begin{pmatrix} M_{11} & \cdots & M_{1d} \\ M_{21} & \cdots & M_{2d} \\ \vdots & \ddots & \vdots \\ M_{d'1} & \cdots & M_{d'd} \end{pmatrix}.$$

[2]By multiplying (4.10) from the *right* by M and following analogous steps as above, one can show that $M^\dagger M = cI$, so that the matrix $c^{-1}M^\dagger$ is both the left and right inverse of M.

We can make a $d' \times d'$ matrix N from M by adding $d' - d$ columns of zeros:

$$N = \begin{pmatrix} M_{11} & \cdots & M_{1d} & 0 & \cdots & 0 \\ M_{21} & \cdots & M_{2d} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ M_{d'1} & \cdots & M_{d'd} & 0 & \cdots & 0 \end{pmatrix} \equiv (M, 0)\,.$$

Taking the adjoint of this matrix yields

$$N^\dagger = \begin{pmatrix} M_{11} & M_{21}^* & \cdots & M_{d'1}^* \\ M_{12}^* & M_{22}^* & \cdots & M_{d'2}^* \\ \vdots & \vdots & \ddots & \vdots \\ M_{1d}^* & M_{2d}^* & \cdots & M_{d'd}^* \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} M^\dagger \\ 0 \end{pmatrix}.$$

Note that this construction maintains the product $MM^\dagger$:

$$NN^\dagger = (M, 0)\begin{pmatrix} M^\dagger \\ 0 \end{pmatrix} = MM^\dagger = cI\,.$$

The determinant of N is clearly zero. Thus,

$$\det(NN^\dagger) = \det(N)\det(N^\dagger) = c^{d'} = 0$$

so $c = 0$, which means that $MM^\dagger = 0$. Proceeding as in Case I, we conclude that this implies $M = 0$. This completes the second part of the proof. ∎

4.3 The Great Orthogonality Theorem

Schur's lemmas provide restrictions on the form of matrices which commute with all of the matrices of irreducible representations. But the

group property enables the construction of many matrices which satisfy the relations in Schur's First and Second Lemmas. The interplay between these two facts provides the basis for proving the Great Orthogonality Theorem. The statement of this theorem is as follows:

Theorem 4.3 (Great Orthogonality Theorem). Let $\{A_1, A_2, \ldots, A_{|G|}\}$ and $\{A'_1, A'_2, \ldots, A'_{|G|}\}$ be two inequivalent irreducible representations of a group G with elements $\{g_1, g_2, \ldots, g_{|G|}\}$ and which have dimensionalities d and d', respectively. The matrices A_α and A'_α in the two representations correspond to the element g_α in G. Then

$$\sum_\alpha (A_\alpha)^*_{ij}(A'_\alpha)_{i'j'} = 0$$

for all matrix elements. For the elements of a single unitary irreducible representation, we have

$$\sum_\alpha (A_\alpha)^*_{ij}(A_\alpha)_{i'j'} = \frac{|G|}{d}\delta_{i,i'}\delta_{j,j'}\,,$$

where d is the dimension of the representation.

Proof. Consider the matrix

$$M = \sum_\alpha A'_\alpha X A_\alpha^{-1}\,, \tag{4.12}$$

where X is an arbitrary matrix with d' rows and d columns, so that M is a $d' \times d'$ matrix. We will show that for any matrix X, M satisfies a commutation relation of the type discussed in Schur's Lemmas.

We now multiply M from the left by the matrix A'_β corresponding to some matrix in the the "primed" representation:

$$\begin{aligned} A'_\beta M &= \sum_\alpha A'_\beta A'_\alpha X A_\alpha^{-1} \\ &= \sum_\alpha A'_\beta A'_\alpha X A_\alpha^{-1} A_\beta^{-1} A_\beta \\ &= \sum_\alpha A'_\beta A'_\alpha X (A_\beta A_\alpha)^{-1} A_\beta\,. \end{aligned} \tag{4.13}$$

Since the A_α and A'_α form representations of G, the products $A_\alpha A_\beta$ and $A'_\alpha A'_\beta$ yield matrices A_γ and A'_γ, respectively, both corresponding to the same element in G because representations preserve the group composition rule. Hence, by the Rearrangement Theorem (Theorem 2.1), we can write the summation over α on the right-hand side of this equation as

$$\sum_\alpha A'_\beta A'_\alpha X(A_\beta A_\alpha)^{-1} = \sum_\gamma A'_\gamma X A_\gamma^{-1} = M\,.$$

Substituting this result into (4.13) yields

$$A'_\beta M = M A_\beta\,. \tag{4.14}$$

Depending on the nature of the two representations, this is precisely the situation addressed by Schur's First and Second Lemmas. We consider the cases of equivalent and inequivalent representations separately.

Case I. $d \neq d'$ or, if $d = d'$, the representations are inequivalent (i.e., not related by a similarity transformation). Schur's Second Lemma then implies that M must be the zero matrix, i.e., that each matrix element of M is zero. From the definition (4.12), we see that this requires

$$M_{ii'} = \sum_\alpha \sum_{jj'} (A'_\alpha)_{ij} X_{jj'} (A_\alpha^{-1})_{j'i'} = 0\,. \tag{4.15}$$

By writing this sum as (note that because all sums are finite, their order can be changed at will)

$$\sum_{jj'} X_{jj'} \Big[\sum_\alpha (A_\alpha)'_{ij} (A_\alpha^{-1})_{j'i'}\Big] = 0\,, \tag{4.16}$$

we see that, since X is arbitrary, each of its entries may be varied arbitrarily and independently without affecting the vanishing of the sum. The only way to ensure this is to require that the coefficients of the $X_{jj'}$ vanish:

$$\sum_\alpha (A'_\alpha)_{ij} (A_\alpha^{-1})_{j'i'} = 0\,.$$

For unitary representations, $(A_\alpha^{-1})_{j'i'} = (A_\alpha)^*_{i'j'}$, so this equation reduces to

$$\sum_\alpha (A'_\alpha)_{ij}(A_\alpha)^*_{i'j'} = 0\,,$$

which proves the first part of the theorem.

Case II. $d = d'$ and the representations are equivalent. According to Schur's First Lemma, $M = cI$, so,

$$cI = \sum_\alpha A_\alpha X A_\alpha^{-1}\,. \tag{4.17}$$

Taking the trace of both sides of this equation,

$$\underbrace{\mathrm{tr}(cI)}_{cd} = \mathrm{tr}\Big(\sum_\alpha A_\alpha X A_\alpha^{-1}\Big) = \sum_\alpha \mathrm{tr}(A_\alpha X A_\alpha^{-1}) = \underbrace{\sum_\alpha \mathrm{tr}(X)}_{|G|\,\mathrm{tr}(X)}\,,$$

yields an expression for c:

$$c = \frac{|G|}{d}\,\mathrm{tr}(X)\,.$$

Substituting this into Eq. (4.17) and expressing the resulting equation in terms of matrix elements, yields

$$\sum_{jj'} X_{jj'}\Big[\sum_\alpha (A_\alpha)_{ij}(A_\alpha^{-1})_{j'i'}\Big] = \frac{|G|}{d}\delta_{i,i'}\sum_j X_{jj}\,,$$

or, after a simple rearrangement,

$$\sum_{jj'} X_{jj'}\Big[\sum_\alpha (A_\alpha)_{ij}(A_\alpha^{-1})_{j'i'} - \frac{|G|}{d}\delta_{i,i'}\delta_{j,j'}\Big] = 0\,.$$

This equation must remain valid under any independent variation of the matrix elements of X. Thus, we must require that the coefficient of $X_{jj'}$ vanishes identically:

$$\sum_\alpha (A_\alpha)_{ij}(A_\alpha^{-1})_{j'i'} = \frac{|G|}{d}\delta_{i,i'}\delta_{j,j'}\,.$$

Since the representation is unitary, this is equivalent to

$$\sum_\alpha (A_\alpha)_{ij}(A_\alpha)^*_{i'j'} = \frac{|G|}{d}\delta_{i,i'}\delta_{j,j'}\,.$$

This proves the second part of the theorem. ∎

4.4 Some Immediate Consequences of the Great Orthogonality Theorem

The Great Orthogonality Theorem establishes a relation between matrix elements of the irreducible representations of a group. Suppose we denote the αth matrix in the kth irreducible representation by A^k_α and the (i,j)th element of this matrix by $(A^k_\alpha)_{ij}$. We can then combine the two statements of the Great Orthogonality Theorem as

$$\sum_\alpha (A^k_\alpha)_{ij}(A^{k'}_\alpha)^*_{i'j'} = \frac{|G|}{d}\delta_{i,i'}\delta_{j,j'}\delta_{k,k'} \tag{4.18}$$

This expression helps us to understand the motivation for the name "Orthogonality Theorem" by inviting us to consider the matrix elements of irreducible representations as entries in $|G|$-component vectors, i.e., vectors in a space of dimensionality $|G|$:

$$\boldsymbol{V}^k_{ij} = \Big[(A^k_1)_{ij}, (A^k_2)_{ij}, \ldots, (A^k_{|G|})_{ij}\Big]$$

According to the statement of the Great Orthogonality Theorem, two such vectors are orthogonal if they differ in any one of the indices i, j, or k, since (4.18) requires that

$$\boldsymbol{V}^k_{ij} \cdot \boldsymbol{V}^{k'}_{i'j'} = \frac{|G|}{d}\delta_{i,i'}\delta_{j,j'}\delta_{k,k'}$$

But, in a $|G|$-dimensional space there are at most $|G|$ mutually orthogonal vectors. To see the consequences of this, suppose we have irreducible representations of dimensionalities $d_1, d_2, \ldots$, where the d_k

are positive integers. For the k representations, there are d_k choices for each of i and j, i.e., there are d_k^2 matrix elements in each matrix of the representation. Summing over all irreducible representations, we obtain the inequality

$$\sum_k d_k^2 \leq |G| \tag{4.19}$$

Thus, the order of the group acts as an upper bound both for the number and the dimensionalities of the irreducible representations. In particular, a finite group can have only a *finite* number of irreducible representations. We will see later that the *equality* in (4.19) always holds.

Example 4.1. For the group S_3, we have that $|G| = 6$ and we have already identified two one-dimensional irreducible representations and one two-dimensional irreducible representation (Example 3.2). Thus, using (4.19), we have

$$\sum_k d_k^2 = 1^2 + 1^2 + 2^2 = 6$$

so the Great Orthogonality Theorem tells us that there are no additional distinct irreducible representations.

For the two element group, we have found two one-dimensional representations, $\{1, 1\}$ and $\{1, -1\}$ (Example 3.3). According to the inequality in (4.19),

$$\sum_k d_k^2 = 1 + 1 = 2$$

so these are the only two irreducible representations of this group. ∎

4.5 Summary

The central result of this chapter is the statement and proof of the Great Orthogonality Theorem. Essentially all of the applications in the next several chapters are consequences of this theorem. The important advance provided this theorem is that it provides an orthogonality

relation between the entries of the matrices of the irreducible representations of a group. While this can be used to test whether a given representation is reducible or irreducible (Problem Set 6), its main role will be in a somewhat "reduced" form, such as that used in Sec. 4.4 to place bounds on the number of irreducible representations of a finite group. One of the most important aspects of the Great Orthogonality Theorem for applications to physical problems is in the construction of "character tables," i.e., tables of traces of matrices of an irreducible representation. This is taken up in the next chapter.

Chapter 5

Characters and Character Tables

In great mathematics there is a very high degree of unexpectedness, combined with inevitability and economy.

—Godfrey H. Hardy[1]

In the preceding chapter, we proved the Great Orthogonality Theorem, which is a statement about the orthogonality between the matrix elements corresponding to different irreducible representations of a group. For many applications of group theory, however, the full matrix representations of a group are not required, but only the traces within classes of group elements—called "characters." A typical application involves determining whether a given representation is reducible or irreducible and, if it is reducible, to identify the irreducible representations contained within that representation.

In this chapter, we develop the mathematical machinery that is used to assemble the characters of the irreducible representations of a group in what are called "character tables." The compilation of character tables requires two types of input: the order of the group and the number of classes it contains. These quantities provide stringent restrictions

[1]G.H. Hardy, *A Mathematician's Apology* (Cambridge University Press, London, 1941)

on the number of irreducible representations and their dimensionalities. Moreover, orthogonality relations derived from the Great Orthogonality Theorem will be shown to provide constraints on characters of different irreducible representations, which considerably simplifies the construction of character tables.

5.1 Orthogonality Relations

The Great Orthogonality Theorem,

$$\sum_\alpha (A_\alpha^k)_{ij}(A_\alpha^{k'})^*_{i'j'} = \frac{|G|}{d_k}\delta_{i,i'}\delta_{j,j'}\delta_{k,k'} \tag{5.1}$$

is a relationship between the matrix elements of the irreducible representations of a group G. In this section, we show how this statement can be manipulated into an expression solely in terms of the traces of the matrices in these representations. This will open the way to establishing a sum rule between the number of irreducible representations and the number of classes in a group.

We begin by setting $j = i$ and $j' = i'$ in (5.1),

$$\sum_\alpha (A_\alpha^k)_{ii}(A_\alpha^{k'})^*_{i'i'} = \frac{|G|}{d_k}\delta_{i,i'}\delta_{k,k'}\,, \tag{5.2}$$

where we have used the fact that $\delta_{i,i'}\delta_{i,i'} = \delta_{i,i'}$. Summing over i and i' on the left-hand side of this equation yields

$$\begin{aligned}\sum_{i,i'}\sum_\alpha (A_\alpha^k)_{ii}(A_\alpha^{k'})_{i'i'} &= \sum_\alpha \underbrace{\left[\sum_i (A_\alpha^k)_{ii}\right]}_{\mathrm{tr}(A_\alpha^k)}\underbrace{\left[\sum_{i'} (A_\alpha^{k'})^*_{i'i'}\right]}_{\mathrm{tr}(A_\alpha^{k'})^*} \\ &= \sum_\alpha \mathrm{tr}(A_\alpha^k)\mathrm{tr}(A_\alpha^{k'})^*\,,\end{aligned}$$

and, by summing over i and i' on the right-hand side of (5.2), we obtain

$$\frac{|G|}{d_k}\delta_{k,k'}\sum_i\sum_{i'}\delta_{i,i'} = \frac{|G|}{d_k}\delta_{k,k'}\underbrace{\sum_i 1}_{d_k} = |G|\delta_{k,k'}\,.$$

We have thereby reduced the Great Orthogonality Theorem to

$$\sum_\alpha \mathrm{tr}(A_\alpha^k)\mathrm{tr}(A_\alpha^{k'*}) = |G|\delta_{k,k'} \,. \tag{5.3}$$

This expression can be written in a more useful form by observing that matrices corresponding to elements in the same conjugacy class have the same trace. To see this, recall the definition in Section 2.6 of the conjugacy of two elements a and b group G. There must be an element g in G such that $a = gbg^{-1}$. Any representation $\{A_\alpha\}$, reducible or irreducible, must preserve this relation:

$$A_a = A_g A_b A_{g^{-1}} \,.$$

This representation must also have the property that $A_{g^{-1}} = A_g^{-1}$. Thus (Problem 2, Problem Set 4),

$$\mathrm{tr}(A_a) = \mathrm{tr}(A_g A_b A_g^{-1}) = \mathrm{tr}(A_g^{-1} A_g A_b) = \mathrm{tr}(A_b) \,.$$

We can now introduce the notation χ_α^k for the trace corresponding to *all* of the elements of the αth class of the kth irreducible representation. This is called the **character** of the class. If there are n_α elements in this class, then we can write the relation (5.3) in terms of characters as a sum over conjugacy classes

$$\sum_{\alpha=1}^{\mathcal{C}} n_\alpha \chi_\alpha^k \chi_\alpha^{k'*} = |G|\delta_{k,k'} \,, \tag{5.4}$$

where $\mathcal{C}$ is the number of conjugacy classes. In arriving at this relation, we have proven the following theorem:

Theorem 5.1 (Orthogonality Theorem for Characters). The characters of the irreducible representations of a group obey the relation

$$\sum_\alpha n_\alpha \chi_\alpha^k \chi_\alpha^{k'*} = |G|\delta_{k,k'} \,.$$

This orthogonality theorem can be used to deduce a relationship between the number classes of a group and the number of irreducible representations. By rearranging (5.4) as

$$\sum_\alpha \left[\left(\frac{n_\alpha}{|G|}\right)^{1/2} \chi_\alpha^k\right]\left[\left(\frac{n_\alpha}{|G|}\right)^{1/2} \chi_\alpha^{k'*}\right] = \delta_{k,k'}$$

and introducing the vectors

$$\widetilde{\boldsymbol{\chi}}^k = |G|^{-1/2}(\sqrt{n_1}\chi_1^k, \sqrt{n_2}\chi_2^k, \ldots, \sqrt{n_\mathcal{C}}\chi_\mathcal{C}^k)\,,$$

we can write the orthogonality relation for characters as

$$\widetilde{\boldsymbol{\chi}}^k \cdot \widetilde{\boldsymbol{\chi}}^{k'} = \delta_{k,k'}\,.$$

The $\widetilde{\boldsymbol{\chi}}^k$ reside in a space whose dimension is the number of classes $\mathcal{C}$ in the group. Thus, the maximum number of a set of mutually orthogonal vectors in this space is $\mathcal{C}$. But these vectors are labelled by an index k corresponding to the irreducible representations of the group. Hence, *the number of irreducible representations must be less than or equal to the number of classes.*

It is also possible[2] to obtain an orthogonality relation with the roles of the irreducible representations and classes reversed in comparison to that in Theorem 5.1:

$$\sum_k \chi_\alpha^k \chi_\beta^{k*} = \frac{|G|}{n_\alpha}\delta_{\alpha,\beta}\,. \tag{5.5}$$

By following analogous reasoning as above, we can deduce that this orthogonality relation implies that *the number of irreducible representations must be greater than or equal to the number of classes.* Combined with the statement of Theorem 5.1, we have the following theorem:

Theorem 5.2. The number of irreducible representations of a group is equal to the number of conjugacy classes of that group.

Example 5.1. For Abelian subgroups each element is in a class by itself (Problem 6, Problem Set 3). Thus, the number of classes is equal to the order of the group, so, according to Theorem 5.2, the number of irreducible representations must also equal the order of the group. When combined with the restriction imposed by Eqn. (4.19), which we can now write as

$$\sum_{k=1}^{|G|} d_k^2 = |G|\,,$$

[2] M. Hamermesh, *Group Theory and its Application to Physical Problems* (Dover, 1989, New York) pp. 106–110.

we have an alternative way (cf. Problem 4, Problem Set 5) of seeing that all of the the irreducible representations of an Abelian group are one-dimensional, i.e., $d_k = 1$, for $k = 1, 2, \ldots, |G|$. ∎

Example 5.2. For the group S_3, there are three classes: $\{e\}$, $\{a, b, c\}$, and $\{d, f\}$ (Example 2.9). Thus, there are three irreducible representations which, as we have seen, consist of two one-dimensional representations and one two-dimensional representation. ∎

5.2 The Decomposition Theorem

One of the main uses of characters is in the decomposition of a given reducible representation into its constituent irreducible representations. The procedure by which this is accomplished is analogous to projecting a vector onto a set of complete orthogonal basis vectors. The theorem which provides the foundation for carrying this out with characters is the following:

Theorem 5.3 (Decomposition Theorem). The character χ_α for the αth class of any representation can be written uniquely in terms of the corresponding characters of the irreducible representations of the group as

$$\chi_\alpha = \sum_k a_k \chi_\alpha^k ,$$

where

$$a_k = \frac{1}{|G|} \sum_\alpha n_\alpha \chi_\alpha^{k*} \chi_\alpha .$$

Proof. For a reducible representation, the same similarity transformation brings all of the matrices into the same block-diagonal form. In this form, the matrix A_α can be written as the direct sum of matrices A_j^k of irreducible representations:

$$A_\alpha = A_\alpha^{k_1} \oplus A_\alpha^{k_2} \oplus \cdots \oplus A_\alpha^{k_n} ,$$

where $\alpha = 1, 2, \ldots, |G|$ and $k_1, k_2, \ldots k_n$ label irreducible representations. Given this, and the fact that similarity transformations leave the trace invariant, we can write the character χ_i of this reducible representation corresponding to the ith class as

$$\chi_\alpha = \sum_k a_k \chi_\alpha^k \,, \tag{5.6}$$

where the a_k must be *nonnegative integers*. We now multiply both sides of this equation by $n_\alpha \chi_\alpha^{k'*}$, sum over α, and use the orthogonality relation (5.4):

$$\sum_\alpha n_\alpha \chi_\alpha^{k'*} \chi_\alpha = \sum_k a_k \underbrace{\sum_\alpha n_\alpha \chi_\alpha^k \chi_\alpha^{k'*}}_{|G|\delta_{k,k'}} = |G| a_{k'}$$

Thus,

$$a_{k'} = \frac{1}{|G|} \sum_\alpha n_\alpha \chi_\alpha^{k'*} \chi_\alpha \,, \tag{5.7}$$

so $a_{k'}$ is the projection of the reducible representation onto the k'th irreducible representation. Note that, because the number of irreducible representations equals the number of classes, the orthogonal vectors of characters span the space whose dimensionality is the number of classes, so this decomposition is unique. ∎

The Decomposition Theorem reduces the task of determining the irreducible representations contained within a reducible representation to one of vector algebra. Unless a particular application requires the matrix forms of the representations, there is no need to block-diagonalize a representation to identify its irreducible components.

We can follow a procedure analogous to that used to prove the Decomposition Theorem to derive a simple criterion to identify whether a representation is reducible or irreducible. We begin with the decomposition (5.6) and take its complex conjugate:

$$\chi_\alpha^* = \sum_{k'} a_{k'} \chi_\alpha^{k'*} \,, \tag{5.8}$$

where we have used the fact that the a_k are integers, so $a_k^* = a_k$. We now take the product of (5.6) and (5.8), multiply by n_α, sum over α, and invoke (5.4):

$$\sum_\alpha n_\alpha \chi_\alpha \chi_\alpha^* = \sum_{k,k'} a_k a_{k'} \underbrace{\sum_i n_\alpha \chi_\alpha^k \chi_\alpha^{k'*}}_{|G|\delta_{k,k'}} = |G| \sum_k a_k^2 \,.$$

Thus,

$$\sum_\alpha n_\alpha |\chi_\alpha^2| = |G| \sum_k a_k^2 \,. \tag{5.9}$$

If the representation in question is irreducible, then all of the a_k are zero, except for the one corresponding to that irreducible representation, which is equal to unity. If the representation is reducible, then there will be at least two of the a_k which are positive integers. We can summarize these observations with a simple criterion for reducibility. If the representation is irreducible, then

$$\sum_\alpha n_\alpha |\chi_\alpha|^2 = |G| \,, \tag{5.10}$$

and if the representation is reducible,

$$\sum_\alpha n_\alpha |\chi_\alpha|^2 > |G| \,. \tag{5.11}$$

Example 5.3. Consider the following representation of S_3:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = \tfrac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad b = \tfrac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix},$$

$$c = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad d = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \qquad f = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}.$$

There are three classes of this group, $\{e\}$, $\{a, b, c\}$, and $\{d, f\}$, so we have $n_1 = 1$, $n_2 = 3$, and $n_3 = 2$, respectively. The corresponding characters are

$$\chi_1 = 2, \qquad \chi_2 = 0, \qquad \chi_3 = -1 \,.$$

Forming the sum in (5.9), we obtain

$$\sum_{\alpha=1}^{3} n_\alpha |\chi_\alpha|^2 = (1\times 4) + (3\times 0) + (2\times 1) = 6\,,$$

which is equal to the order of the group. Therefore, this representation is *irreducible*, as we have already demonstrated in Example 3.4 and in Problem 1, Problem Set 6.

Example 5.4. Another representation of S_3 is

$$e = d = f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = b = c = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}.$$

The characters corresponding to the three classes are now

$$\chi_1 = 2, \qquad \chi_2 = 0, \qquad \chi_3 = 2\,.$$

Forming the sum in (5.9), we find

$$\sum_{i=1}^{3} n_1 |\chi_i|^2 = (1\times 4) + (3\times 0) + (2\times 4) = 12\,,$$

which is *greater* than the order of the group, so this representation is *reducible* (cf. Problem 2, Problem Set 6). To determine the irreducible constituents of this representation, we use the decomposition theorem. There are three irreducible representations of S_3: the one-dimensional identical representation, with characters

$$\chi_1^1 = 1, \qquad \chi_2^1 = 1, \qquad \chi_3^1 = 1\,,$$

the one-dimensional "parity" representation, with characters

$$\chi_1^2 = 1, \qquad \chi_2^2 = -1, \qquad \chi_3^2 = 1\,,$$

and the two-dimensional "coordinate" representation discussed above in Example 5.3, with characters

$$\chi_1^3 = 2, \qquad \chi_2^3 = 0, \qquad \chi_3^3 = -1\,.$$

We now calculate the a_k using the expression in Equation (5.7). These determine the "projections" of the characters of the reducible representation onto the characters of the irreducible representation. We obtain

$$a_1 = \tfrac{1}{6}\Big[(1 \times 1 \times 2) + (3 \times 1 \times 0) + (2 \times 1 \times 2)\Big] = 1\,,$$

$$a_2 = \tfrac{1}{6}\Big[(1 \times 1 \times 2) + (3 \times -1 \times 0) + (2 \times 1 \times 2)\Big] = 1\,,$$

$$a_3 = \tfrac{1}{6}\Big[(1 \times 2 \times 2) + (3 \times 0 \times 0) + (2 \times -1 \times 2)\Big] = 0\,.$$

Thus, this reducible representation is composed of the identical representation and the "parity" representation, with no contribution from the "coordinate" representation. The block-diagonal form of this representation is, therefore,

$$e = d = f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = b = c = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is the result obtained in Problem 5, Problem Set 5 by applying matrix methods. ∎

5.3 The Regular Representation

Our construction of irreducible representations has thus far proceeded in an essentially *ad hoc* fashion, relying in large part on physical arguments. We have not yet developed a systematic procedure for constructing all of the irreducible representations of a group. In this section, we introduce a method, based on what is called the "regular" representation, which enables us to accomplish this. However, our purpose for introducing such a methodology is not the determination of irreducible representations as such, since even for relatively simple groups, the approach we describe would present a computationally demanding process, but as a theoretical tool for proving a theorem. Moreover, we will find that, for applications of group theory to quantum mechanics, the irreducible representations of the group of operations that leave Hamiltonian invariant will emerge naturally without having to rely on any auxiliary constructions.

The **regular representation** is a reducible representation that is obtained directly from the multiplication table of a group. As we will show below, this representation contains every irreducible representation of a group at least once. The construction of the regular representation is based on arranging the multiplication table of a group so that the unit element appears along the main diagonal of the table. Within such an arrangement the columns (or rows) of the table are labelled by the group elements, arranged in any order, and the corresponding order of the inverses labels the rows (or columns).

As an example, consider the multiplication table for S_3 (Section 2.4) arranged in the way just described:

	e	a	b	c	d	f
$e = e^{-1}$	e	a	b	c	d	f
$a = a^{-1}$	a	e	d	f	b	c
$b = b^{-1}$	b	f	e	d	c	a
$c = c^{-1}$	c	d	f	e	a	b
$f = d^{-1}$	f	b	c	a	e	d
$d = f^{-1}$	d	c	a	b	f	e

The matrices of the regular representation are obtained by regarding the multiplication table as an $|G| \times |G|$ array from which the matrix representation for each group element is assembled by putting a '1' where that element appears in the multiplication table and zero elsewhere. For example, the matrices corresponding to the unit e and the element a are

$$e \to \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad a \to \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

with analogous matrices for the other group elements.

Our first task is to show that the regular 'representation' is indeed a representation of the group. First of all, it is clear that the mapping we have described is one-to-one. For any two elements g_1 and g_2 of this group, we denote the matrices in the regular representation that correspond to these elements as $A_{\rm reg}(g_1)$ and $A_{\rm reg}(g_2)$. Thus, to show that these matrices form a representation of S_3, we need to verify that

$$A_{\rm reg}(g_1g_2) = A_{\rm reg}(g_1)A_{\rm reg}(g_2)\,,$$

i.e., that the multiplication table is preserved by this representation. We consider this relation expressed in terms of matrix elements:

$$\Big[A_{\rm reg}(g_1g_2)\Big]_{ij} = \sum_k \Big[A_{\rm reg}(g_1)\Big]_{ik}\Big[A_{\rm reg}(g_2)\Big]_{kj}\,. \tag{5.12}$$

From the way the regular representation has been constructed, the ith row index of these matrix elements can be labelled the inverse of the ith group element g_i^{-1} and the jth column can be labelled by the jth group element g_j:

$$\Big[A_{\rm reg}(g_1g_2)\Big]_{ij} = \Big[A_{\rm reg}(g_1g_2)\Big]_{g_i^{-1},g_j} = \begin{cases} 1, & \text{if } g_i^{-1}g_j = g_1g_2; \\ 0; & \text{otherwise} \end{cases}$$

$$\Big[A_{\rm reg}(g_1)\Big]_{ik} = \Big[A_{\rm reg}(g_1)\Big]_{g_i^{-1},g_k} = \begin{cases} 1, & \text{if } g_i^{-1}g_k = g_1; \\ 0; & \text{otherwise} \end{cases}$$

$$\Big[A_{\rm reg}(g_2)\Big]_{kj} = \Big[A_{\rm reg}(g_2)\Big]_{g_k^{-1},g_j} = \begin{cases} 1, & \text{if } g_k^{-1}g_j = g_2; \\ 0; & \text{otherwise} \end{cases}$$

Therefore, in the sum over k in (5.12), we have nonzero entries only when

$$g_1g_2 = (g_i^{-1}g_k)(g_k^{-1}g_j) = g_i^{-1}g_j\,,$$

which gives precisely the nonzero matrix elements of $A_{\rm reg}(g_1g_2)$. Hence, the matrices $A_{\rm reg}(g_1)$ preserve the group multiplication table and thereby form a faithful representation of the group.

Our main purpose in introducing the regular representation is to prove the following theorem:

Theorem 5.4. The dimensionalities d_k of the irreducible representations of a group are related to the order $|G|$ of the group by

$$\sum_k d_k^2 = |G|\,.$$

This theorem shows that the inequality (4.19), which was deduced directly from the Great Orthogonality Theorem is, in fact, an equality.

Proof. We first show, using Eqn. (5.9), that the regular representation is reducible. To evaluate the sums on the left-hand side of this equation, we note that, from the construction of the regular representation, the characters $\chi_{\mathrm{reg},i}$ vanish for every class except for that corresponding to the unit element. Denoting this character by $\chi_{\mathrm{reg},e}$, we see that its value must be equal to the order of the group:

$$\chi_{\mathrm{reg},e} = |G|\,.$$

Thus,

$$\sum_\alpha n_\alpha|\chi_\alpha|^2 = \chi_{\mathrm{reg},e}^2 = |G|^2\,,$$

which, for $|G| > 1$ is *greater* than $|G|$. Thus, for groups other than the single-element group $\{e\}$, the regular representation is reducible.

We will now use the Decomposition Theorem to identify the irreducible constituents of the regular representation. Thus, the characters $\chi_{\mathrm{reg},\alpha}$ for the αth class in the regular representation can be written as

$$\chi_{\mathrm{reg},\alpha} = \sum_k a_k\chi_\alpha^k\,.$$

According to the Decomposition Theorem, the a_k are given by

$$a_k = \frac{1}{|G|}\sum_\alpha n_\alpha\chi_\alpha^{k*}\chi_{\mathrm{reg},\alpha}\,.$$

We again use the fact that $\chi_{\text{reg},e} = |G|$, with all other characters vanishing. The corresponding value of χ_e^k is determined by taking the trace of the identity matrix whose dimensionality is that of the kth irreducible representation: $\chi_e^k = d_k$. Therefore, the Decomposition Theorem yields

$$a_k = \frac{1}{|G|} \times d_k \times |G| = d_k \,,$$

i.e., the kth irreducible representation appears d_k times in the regular representation: each one-dimensional irreducible representation appears once, each two-dimensional irreducible representation appears twice, and so on. Since the dimensionality of the regular representation is $|G|$, and since a_k is the number of times the kth irreducible representation appears in the regular representation, we have the constraint

$$\sum_k a_k d_k = |G| \,,$$

i.e.,

$$\sum_k d_k^2 = |G| \,.$$

∎

This sum rule, and that equating the number of classes to the number of irreducible representations (Theorem 5.2), relate a property of the abstract group (its order and the number of classes) to a property of the irreducible representations (their number and dimensionality). The application of these rules and the orthogonality theorems for characters is the basis for constructing character tables. This is described in the next section.

5.4 Character Tables

Character tables are central to many applications of group theory to physical problems, especially those involving the decomposition of reducible representations into their irreducible components. Many textbooks on group theory contain compilations of character tables for the

most common groups. In this section, we will describe the construction of character tables for S_3. We will utilize two types of information: sum rules for the number and dimensionalities of the irreducible representations, and orthogonality relations for the characters. Additionally, the group multiplication table can be used to establish relationships for one-dimensional representations. By convention, characters tables are displayed with the columns labelled by the classes and the rows by the irreducible representations.

The first step in the construction of this character table is to note that, since $|S_3| = 6$ and there are three classes (Example 2.9), there are 3 irreducible representations whose dimensionalities must satisfy

$$d_1^2 + d_2^2 + d_3^2 = 6\,.$$

The unique solution of this equation (with only positive integers) is $d_1 = 1$, $d_2 = 1$, and $d_3 = 2$, so there are two one-dimensional irreducible representations and one two-dimensional irreducible representation.

In the character table for any group, several entries can be made immediately. The identical representation, where all elements are equal to unity, is always a one-dimensional irreducible representation. Similarly, the characters corresponding to the unit element are equal to the dimensionality of that representation, since they are calculated from the trace of the identity matrix with that dimensionality. Thus, denoting by α, β, γ, and δ quantities that are to be determined, the character table for S_3 is:

S_3	$\{e\}$	$\{a,b,c\}$	$\{d,f\}$
Γ_1	1	1	1
Γ_2	1	α	β
Γ_3	2	γ	δ

where the Γ_i are a standard label for the irreducible representations.

The remaining entries are determined from the orthogonality relations for characters and, for one-dimensional irreducible representations, from the multiplication table of the group. The orthogonality relation in Theorem 5.1, which is an orthogonality relation for the *rows* of a character table, yield

$$1 + 3\alpha + 2\beta = 0\,, \tag{5.13}$$

$$1 + 3\alpha^2 + 2\beta^2 = 6\,. \tag{5.14}$$

The group multiplication table requires that

$$a^2 = e, \qquad b^2 = e, \qquad c^2 = e, \qquad d^2 = f\,.$$

Since the one-dimensional representations must obey the multiplication table, these products imply that

$$\alpha^2 = 1, \qquad \beta^2 = \beta\,.$$

Substituting these relations into (5.14), yields $4 + 2\beta = 6$, i.e.,

$$\beta = 1$$

Upon substitution of this value into (5.13), we obtain $3 + 3\alpha = 0$, i.e.,

$$\alpha = -1$$

From the orthogonality relation (5.5), which is an orthogonality relation between the *columns* of a character table, we obtain

$$1 + \alpha + 2\gamma = 0$$

$$1 + \beta + 2\delta = 0$$

Substituting the values obtained for α and β into these equations yields

$$\gamma = 0, \qquad \delta = -1$$

The complete character table for S_3 is therefore given by

S_3	$\{e\}$	$\{a,b,c\}$	$\{d,f\}$
Γ_1	1	1	1
Γ_2	1	-1	1
Γ_3	2	0	-1

When character tables are compiled for the most common groups, a notation is used which reflects the fact that the group elements correspond to transformations on physical objects. The notation for the classes of S_3 are as follows:

- $\{e\} \to E$. The identity.

- $\{a, b, c\} \to 3\sigma_v$. Reflection through *vertical* planes, where 'vertical' refers to the fact that these planes contain the axis of highest rotational symmetry, in this case, the z-axis. The '3' refers to there being three elements in this class.

- $\{d, f\} \to 2C_3$. Rotation by $\frac{2}{3}\pi$ radians, with the '2' again referring to the there being two elements in this class. The notation C_3^2 is for rotations by $\frac{4}{3}\pi$ radians, so the 'class' notation is meant only to indicate the *type* of operation. In general, C_n refers to rotations through $2\pi/n$ radians.

Several notations are used for irreducible representations. One of the most common is to use A for one-dimensional representations, E for two-dimensional representations, and T for three-dimensional representations, with subscripts used to distinguish multiple occurrences of irreducible representations of the same dimensionality. The notation Γ is often used to indicate a generic (usually irreducible) representation, with subscripts and superscripts employed to distinguish between different representations. With the first of these conventions, the character table for S_3, which is known as the group C_{3v} when interpreted as the planar symmetry operations of an equilateral triangle, is

C_{3v}	E	$3\sigma_v$	$2C_3$
A_1	1	1	1
A_2	1	-1	1
E	2	0	-1

∎

5.5 Summary

This chapter has been devoted to characters and character tables. The utility of characters in applications stems from the following:

1. The character is a property of the class of an element.

2. Characters are unaffected by similarity transformations, so equivalent representations—reducible or irreducible—have the same characters.

3. As shown in Equations (5.10) and (5.11), the characters of a representation indicate, through a straightforward calculation, whether that representation is reducible or irreducible.

4. Characters of irreducible representations obey orthogonality theorems which, when interpreted in the context of character tables, correspond to the orthogonality relations of their rows and columns.

5. According to the Decomposition Theorem, once the character table of a group is known, the characters of any representation can be decomposed into its irreducible components.

Chapter 6

Groups and Representations in Quantum Mechanics

The universe is an enormous direct product of representations of symmetry groups.

—Steven Weinberg[1]

This chapter is devoted to applying the mathematical theory of groups and representations which we have developed in the preceding chapters to the quantum mechanical description of physical systems. The power of applying group theory to quantum mechanics is that it provides a framework for making *exact* statements about a physical system with a knowledge only of the symmetry operations which leave its Hamiltonian invariant, the so-called "group of the Hamiltonian." Moreover, when we apply the machinery of groups to quantum mechanics, we find that representations—and irreducible representations in particular—arise quite naturally, as do related concepts such as the importance of unitarity of representations and the connection between the symmetry of a physical system and the degeneracy of its eigenstates. We will follow the general sequence of the discussion in Sections 1.2 and 1.3,

[1]Steven Weinberg, Sheldon Glashow, and Abdus Salam were awarded the 1979 Nobel Prize in Physics for their incorporation of the weak and electromagnetic interactions into a single theory.

beginning with the group of the Hamiltonian, using this to establish the symmetry properties of the eigenfunctions, and concluding with a discussion of selection rules, which demonstrates the power and economy of using character tables. As a demonstration of the usefulness of these constructions, we will prove Bloch's theorem, the fundamental principle behind the properties of wavefunctions in periodic systems such as electrons and phonons (the quanta of lattice vibrations) in single crystals. The application of group theory to selection rules necessitates the introduction of the "direct product" of matrices and groups, though here, too, quantum mechanics provides a motivation for this concept.

6.1 The Group of the Hamiltonian

Recall the definition of a similarity transformation introduced in Section 3.3. Two matrices, or operators, A and B are related by a similarity transformation generated by a matrix (or operator) R if

$$B = RAR^{-1}\,.$$

The quantity B is therefore the expression of A under the transformation R. Consider now a Hamiltonian $\mathcal{H}$ and its transformation by an operation R

$$R\mathcal{H}R^{-1}\,.$$

The Hamiltonian is said to be *invariant* under R if

$$\mathcal{H} = R\mathcal{H}R^{-1}\,, \tag{6.1}$$

or, equivalently,

$$R\mathcal{H} = \mathcal{H}R\,. \tag{6.2}$$

Thus the *order* in which $\mathcal{H}$ and the R are applied is immaterial, so $\mathcal{H}$ and R commute: $[\mathcal{H}, R] = 0$. In this case, R is said to be a **symmetry operation** of the Hamiltonian.

Consider set of all symmetry operations of the Hamiltonian, which we will denote by $\{R_\alpha\}$. We now show that these operations form a

group. To demonstrate closure, we observe that if R_α and R_β are two operations which satisfy (6.1), then

$$R_\alpha \mathcal{H} R_\alpha^{-1} = R_\alpha(R_\beta \mathcal{H} R_\beta^{-1})R_\alpha^{-1} = (R_\alpha R_\beta)\mathcal{H}(R_\alpha R_\beta)^{-1} = \mathcal{H}\,.$$

Thus, the product $R_\alpha R_\beta = R_\gamma$ is also a symmetry operation of the Hamiltonian. Associativity is clearly obeyed since these operations represent transformations of coordinates and other variables of the Hamiltonian.[2] The unit element E corresponds to performing no operation at all and the inverse R_α^{-1} of a symmetry operation R_α is the application of the reverse operation to "undo" the original transformation. Thus, the set $\{R_\alpha\}$ forms a group, called the **group of the Hamiltonian**.

6.2 Eigenfunctions and Representations

There are a number of consequences of the discussion in the preceding section for the representations of the group of the Hamiltonian. Consider an eigenfunction φ of a Hamiltonian $\mathcal{H}$ corresponding to the eigenvalue E:

$$\mathcal{H}\varphi = E\varphi\,.$$

We now apply a symmetry operation R_α to both sides of this equation,

$$R_\alpha \mathcal{H}\varphi = E R_\alpha \varphi\,,$$

and use (6.2) to write

$$R_\alpha \mathcal{H}\varphi = \mathcal{H} R_\alpha \varphi\,.$$

Thus, we have

$$\mathcal{H}(R_\alpha \varphi) = E(R_\alpha \varphi)\,.$$

If the eigenvalue is nondegenerate, then $R_\alpha\varphi$ differs from φ by at most a phase factor:

$$R_\alpha \varphi = \mathrm{e}^{i\phi_\alpha}\varphi\,.$$

[2]The associativity of linear operations is discussed by Wigner in *Group Theory* (Academic, New York, 1959), p. 5.

The application of a second operation R_β then produces

$$R_\beta(R_\alpha\varphi) = \mathrm{e}^{i\phi_\beta}\mathrm{e}^{i\phi_\alpha}\varphi\,. \tag{6.3}$$

The left-hand side of this equation can also be written as

$$(R_\beta R_\alpha)\varphi = \mathrm{e}^{i\phi_{\beta\alpha}}\varphi\,, \tag{6.4}$$

Equating the right-hand sides of Eqs. (6.3) and (6.4), yields

$$\mathrm{e}^{i\phi_{\beta\alpha}} = \mathrm{e}^{i\phi_\beta}\mathrm{e}^{i\phi_\alpha}\,,$$

i.e., these phases preserve the multiplication table of the symmetry operations. Thus, the repeated application of all of the R_α to φ generates a *one-dimensional representation* of the group of the Hamiltonian.

The other case to consider occurs if the application of all of the symmetry operations to φ produces ℓ distinct eigenfunctions. These eigenfunctions are said to be ℓ-fold degenerate. If these are the *only* eigenfunctions which have energy E, this is said to be a **normal degeneracy**. If, however, there are other degenerate eigenfunctions which are not captured by this procedure, this is said to be an **accidental degeneracy**. The term “accidental” refers to the fact that the degeneracy is not due to symmetry. But an “accidental” degeneracy can also occur because a symmetry is “hidden,” i.e., not immediately apparent, so the group of the Hamiltonian is not complete. One well-known example of this is the level degeneracy of the hydrogen atom.

For a normal degeneracy, there are orthonormal eigenfunctions φ_i, $i = 1, 2, \ldots, \ell$ which, upon application of one of the symmetry operations R_α are transformed into linear combinations of one another. Thus, if we denote by $\boldsymbol{\varphi}$ the ℓ-dimensional row vector

$$\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \ldots, \varphi_\ell)\,,$$

we can write

$$R_\alpha\boldsymbol{\varphi} = \boldsymbol{\varphi}\Gamma(R_\alpha)\,,$$

where $\Gamma(R_\alpha)$ is an $\ell\times\ell$ matrix. In terms of components, this equation reads

$$R_\alpha\varphi_i = \sum_{k=1}^{\ell}\varphi_k[\Gamma(R_\alpha)]_{ki} \tag{6.5}$$

The successive application of operations R_α and R_β then yields

$$R_\beta R_\alpha \phi_i = R_\beta \sum_{k=1}^{\ell} \varphi_k [\Gamma(R_\alpha)]_{ki} = \sum_{k=1}^{\ell} (R_\beta \varphi_k)[\Gamma(R_\alpha)]_{ki} \,.$$

The operation $R_\beta \varphi_k$ can be written as in (6.5):

$$R_\beta \varphi_k = \sum_{j=1}^{\ell} \varphi_j [\Gamma(R_\beta)]_{jk} \,.$$

Thus,

$$\begin{aligned} R_\beta R_\alpha \phi_i &= \sum_{k=1}^{\ell} \sum_{j=1}^{\ell} \varphi_j [\Gamma(R_\beta)]_{jk} [\Gamma(R_\alpha)]_{ki} \\ &= \sum_{j=1}^{\ell} \varphi_j \left\{ \sum_{k=1}^{\ell} [\Gamma(R_\beta)]_{jk} [\Gamma(R_\alpha)]_{ki} \right\} . \end{aligned} \tag{6.6}$$

Alternatively, we can write

$$R_\beta R_\alpha \varphi_i = \sum_{j=1}^{\ell} \varphi_j [\Gamma(R_\beta R_\alpha)]_{ji} \,. \tag{6.7}$$

By comparing (6.6) and (6.7) and using the orthonormality of the wavefunctions, we conclude that

$$\Gamma(R_\beta R_\alpha) = \Gamma(R_\beta)\Gamma(R_\alpha) \,,$$

so the $\Gamma(R_i)$ form an ℓ-dimensional representation of the group of the Hamiltonian. Since the eigenfunctions can be made orthonormal, this representation can always be taken to be *unitary* (Problem Set 8). We will now show that this representation is also *irreducible*. We first consider the effect of replacing the φ_i by a linear combination of these functions, $\boldsymbol{\psi} = \boldsymbol{\varphi} U$. Then the effect of operating with R on the $\boldsymbol{\psi}$ is

$$R\boldsymbol{\psi} = R\boldsymbol{\varphi} U = \boldsymbol{\varphi} \Gamma U = \boldsymbol{\psi} U^{-1} \Gamma U \,,$$

i.e., the representation with the transformed wavefunctions is related by a similarity transformation to that with the original eigenfunctions,

i.e., the two representations are *equivalent*. Suppose that this representation is reducible. Then there is a unitary transformation of the φ_j such that there are two or more subsets of the ψ_i that transform only among one another under the symmetry operations of the Hamiltonian. This implies that the application of the R_i to any eigenfunction generates eigenfunctions only in the same subset. The degeneracy of the eigenfunctions in the other subset is therefore *accidental*, in contradiction to our original assertion that the degeneracy is *normal*. Hence, the representation obtained for a normal degeneracy is *irreducible* and the corresponding eigenfunctions are said to *generate*, or form a **basis** for this representation.

We can summarize the results of this section as follows:

- To each eigenvalue of a Hamiltonian there corresponds a unique irreducible representation of the group of that Hamiltonian.
- The degeneracy of an eigenvalue is the dimensionality of this irreducible representation. Thus, the dimensionalities of the irreducible representations of a group are the possible degeneracies of Hamiltonians with that symmetry group.
- Group theory provides "good quantum numbers," i.e., labels corresponding irreducible representations to which eigenfunctions belong.
- Although these statements have been shown for *finite* groups, they are also valid for *continuous* groups.

6.3 Group Theory in Quantum Mechanics

The fact that eigenfunctions corresponding to an ℓ-fold degenerate eigenvalue form a basis for an ℓ-dimensional irreducible representation of the group of the Hamiltonian is one of the fundamental principles behind the application of group theory to quantum mechanics. In this section, we briefly describe the two main types of such applications, namely, where group theory is used to obtain *exact* results, and where it is used in conjunction with perturbation theory to obtain *approximate* results.

6.3.1 Exact Results

One of the most elegant applications of group theory to quantum mechanics involves using the group of the Hamiltonian to determine the (normal) degeneracies of the eigenstates, which are just the dimensions of the irreducible representations. Because such a classification is derived from the symmetry properties of the Hamiltonian, it can be accomplished without having to *solve* the Schrödinger equation. Among the most historically important of such applications is the classification of atomic spectral lines. The atomic Hamiltonian is comprised of the sum of the kinetic energies of the electrons and their Coulomb interactions, so an exact solution is impractical, even for few-electron atoms such as He. Nevertheless, the spherical symmetry of the Hamiltonian enables the identification of the irreducible representations of atomic states from which are derived the angular momentum addition rules and multiplet structures. This will be explored further when we discuss continuous groups. Another exact result is Bloch's theorem, which is the basis for many aspects of condensed matter physics. This theorem uses the translational invariance of perfect periodic crystals to determine the form of the eigenfunctions. As discussed in the next section, Bloch's theorem can be reduced to a statement about the (one-dimensional) irreducible representations and basis functions of cyclic groups.

The lowering of the symmetry of a Hamiltonian by a perturbation can also be examined with group theory. In particular, the question of whether the allowed degeneracies are affected by such a perturbation can be addressed by examining the irreducible representations of the groups of the original and perturbed Hamiltonians. Group theory can address not only whether degeneracies can change (from the irreducible representations of the two groups), but how irreducible representations of the original group are related to those of the perturbed group. Typically, when the symmetry of a system is lowered, the dimensionalities of the irreducible representations can also be lowered, resulting in a "splitting" of the original irreducible representations into lower-dimensional irreducible representations of the group of the perturbed system.

Finally, on a somewhat more practical level, group theory can be used to construct symmetrized linear combinations of basis functions

to diagonalize a Hamiltonian. Examples where this arises is the lowering of the symmetry of a system by a perturbation, where the basis functions are the eigenfunctions of the original Hamiltonian, the bonding within molecules, where the basis functions are localized around the atomic sites within the molecule, and vibrations in molecules and solids, where the basis functions describe the displacements of atoms. These applications are discussed by Tinkham.[3]

6.3.2 Approximate Results

The most common application of group theory in approximate calculations involves the calculation of matrix elements in perturbation theory. A typical example is involved adding to a Hamiltonian $\mathcal{H}_0$ and perturbation $\mathcal{H}'$ due to an electromagnetic field which causes transitions between the eigenstates of the original Hamiltonian. The transition rate W is calculated from first-order time-dependent perturbation theory, with the result known as *Fermi's Golden Rule*:[4]

$$W = \frac{2\pi}{\hbar}\varrho_{if}|(i|\mathcal{H}'|f)|^2 \,,$$

where ϱ_{if} is called the "joint density of states," which is a measure of the number of initial and final states which are available for the excitation, and $(i|\mathcal{H}'|f)$ is a matrix element of $\mathcal{H}'$ between the initial and final states. The application of group theory to this problem, which is the subject of Section 6.6, involves determining when this matrix element vanishes by reasons of symmetry.

6.4 Bloch's Theorem*

Bloch's theorem is of central importance to many aspects of electrons, phonons, and other excitations in crystalline solids. One of the main results of this theorem, namely, the form of the eigenfunctions, can be derived solely from group theory. We will work in one spatial dimension,

[3] M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw–Hill, New York, 1964)

[4] L.I. Schiff, *Quantum Mechanics* 2nd edn (McGraw–Hill, New York, 1955)

but the discussion can be extended easily to higher dimensions. We consider a one-dimensional crystal where the distance between nearest neighbors is a and the number of repeat units is N (a large number for a macroscopic solid). Since this system is finite, it has no translational symmetry. However, by imposing a type of boundary condition known as *periodic*, whereby the Nth unit is identified with the first unit—effectively forming a circle from this solid—we now have N discrete symmetries. The Schrödinger equation for a particle of mass m moving in the periodic potential of this system is

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+V(x)\right]\varphi=E\varphi\,,$$

where $V(x+a)=V(x)$.

6.4.1 The Group of the Hamiltonian

The translation of an eigenfunction by a will be denoted by R_a:

$$R_a\varphi(x)=\varphi(x+a)\,.$$

The basic properties of translations originate with the observation that a translation through na,

$$R_{na}\varphi(x)=\varphi(x+na)\,,$$

can be written as the n-fold product of R_a:

$$R_a^n\varphi(x)=\underbrace{R_aR_a\cdots R_a}_{n\,\text{factors}}\varphi(x)=\varphi(x+na)\,.$$

Moreover, because of the periodic boundary conditions, we identify the Nth unit with the first, so

$$R_a^N=R_0\,,$$

which means that no translation is carried out at all. Thus, the collection of all the translations can be written as the powers of a single element, R_a:

$$\{R_a,R_a^2,\ldots,R_a^N=E\}\,, \tag{6.8}$$

where E is the identity. This shows that the group of the Hamiltonian is a cyclic group of order N. In particular, since cyclic groups are Abelian, there are N one-dimensional irreducible representations of this group, i.e., each eigenvalue is nondegenerate and labelled by one of these irreducible representations.

6.4.2 Character Table and Irreducible Representations

Having identified the algebraic structure of the group of the Hamiltonian, we now construct the character table. Since $R_a^N = E$, and since all irreducible representations are one-dimensional, the character for R_a in each of these representations, $\chi^{(n)}(R_a)$ must obey this product:

$$\left[\chi^{(n)}(R_a)\right]^N = 1\,.$$

The solutions to this equation are the Nth roots of unity (cf. Problem 3, Problem Set 5):

$$\chi^{(n)}(R_a) = \mathrm{e}^{2\pi i n/N}, \qquad n = 0, 1, 2, \ldots N-1\,.$$

The character table is constructed by choosing one of these values for each irreducible representation and then determining the remaining entries from the multiplication table of the group (since each irreducible representation is one-dimensional). The resulting character table is:

	$\{E\}$	$\{R_a\}$	$\{R_a^2\}$	$\cdots$	$\{R_a^{N-1}\}$
Γ_1	1	1	1	$\cdots$	1
Γ_2	1	ω	ω^2	$\cdots$	ω^{N-1}
Γ_3	1	ω^2	ω^4	$\cdots$	$\omega^{2(N-1)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Γ_N	1	ω^{N-1}	ω^{2N-2}	$\cdots$	$\omega^{(N-1)^2}$

where $\omega = \mathrm{e}^{2\pi i/N}$. If we denote the eigenfunction corresponding to the nth irreducible representation by φ_n, then applying R_a yields

$$R_a\varphi_n(x) = \omega^{n-1}\varphi_n(x) = \varphi_n(x+a)\,.$$

Since the characters of this group are pure phases, the moduli of the eigenfunctions are periodic functions:

$$|\varphi_n(x+a)|^2 = |\varphi_n(x)|^2 .$$

Thus, the most general form of the φ_n is

$$\varphi_n(x) = \mathrm{e}^{i\phi_n(x)} u_n(x) , \tag{6.9}$$

where $\phi_n(x)$ is a phase function, which we will determine below, and the u_n have the periodicity of the lattice: $u_n(x+a) = u_n(x)$. By combining this form of the wavefunction with the transformation properties required by the character table, we can write

$$R_a^m \varphi_n(x) = \omega^{m(n-1)} \varphi_n(x) = \omega^{m(n-1)} \mathrm{e}^{i\phi_n(x)} u_n(x) .$$

Alternatively, by applying the same translation operation directly to (6.9) yields

$$R_a^m \varphi_n(x) = \varphi_n(x+ma) = \mathrm{e}^{i\phi_n(x+ma)} u_n(x) .$$

By equating these two ways of writing $R_a^m \varphi_n(x)$, we find that their phase changes must be equal. This, in turn, requires that the phase function satisfies

$$\phi_n(x+ma) = \phi_n(x) + \frac{2\pi m(n-1)}{N} . \tag{6.10}$$

Thus, ϕ_n is a *linear* function of m and, therefore, also of $x+ma$, since ϕ_n is a function of only a single variable:

$$\phi_n(x) = Ax + B ,$$

where A and B are constants to be determined. Upon substitution of this expression into both sides of (6.10),

$$A(x+ma) + B = Ax + B + \frac{2\pi m(n-1)}{N} ,$$

and cancelling common factors, we obtain

$$\phi_n(x) = k_n x + B ,$$

where

$$k_n = \frac{2\pi(n-1)}{Na} = \frac{2\pi(n-1)}{L}$$

and $L = Na$ is the size system. The wavefunction in (6.9) thereby reduces to

$$\varphi_n(x) = \mathrm{e}^{ik_n x} u_n(x)\,,$$

where we have absorbed the constant phase due to B into the definition of $u_n(x)$. This is called a **Bloch function**: a function $u_n(x)$ with the periodicity of the lattice modulated by a plane wave.[5] This is one of the two main results of Bloch's theorem, the other being the existence of energy gaps, which is beyond the scope of the discussion here.

6.5 Direct Products

The direct product provides a way of enlarging the number of elements in a group while retaining the group properties. Direct products occur in several contexts. For example, if a Hamiltonian or Lagrangian contains different types of coordinates, such as spatial coordinates for different particles, or spatial and spin coordinates, then the symmetry operations on the different coordinates commute with each other. If there is a coupling between such degrees of freedom, such as particle interactions or a spin-orbit interaction, then the direct product is required to determine the appropriate irreducible representations of the resulting eigenstates. In this section, we develop the group theory associated with direct products and their representations. We will then apply these concepts to selection rules in the following section.

[5] A related issue which can be addressed by group theory is the nature of the quantity $\hbar k_n$. Although it has units of momentum, it does not represent a true momentum, but is called the "crystal momentum." The true momentum $\hbar k$ labels the irreducible representations of the translation group, which is a continuous group and will be discussed in the next chapter. The discrete translations of a periodic potential form a subgroup of the full translation group, so the corresponding irreducible representations *cannot* be labelled by momentum.

6.5.1 Direct Product of Groups

Consider two groups

$$G_a = \{e, a_2, \ldots, a_{|G_a|}\}, \quad G_b = \{e, b_2, \ldots, b_{|G_b|}\},$$

such that all elements in G_a commute with all elements in G_b:

$$a_i b_j = b_j a_i ,$$

for $i = 1, 2, \ldots, |G_a|$ and $j = 1, 2, \ldots, |G_b|$. We have defined $a_1 = e$ and $b_1 = e$. The **direct product** of G_a and G_b, denoted by $G_a \otimes G_b$, is the set containing all elements $a_i b_j$:

$$G_a \otimes G_b = \{e, a_2, \ldots, a_{|G_a|}, b_2, \ldots, b_{|G_b|}, \ldots, a_i b_j, \ldots\} . \qquad (6.11)$$

As shown in Problem 3 of Problem Set 8, the direct product is a group of order $|G_a||G_b|$.

Example 6.1. Consider the symmetry operations on an equilateral triangle that has a *thickness*, i.e., the triangle has become a “wedge.” Thus, in addition to the original symmetry operations of the planar equilateral triangle, there is now also a reflection plane σ_h. There are now six vertices, which are labelled as in Example 2.1, except that we now distinguish between points which lie above, $\{1^+, 2^+, 3^+\}$, and below, $\{1^-, 2^-, 3^-\}$, the reflection plane. The original six operations do not transform points above and below the reflection plane into one another. The reflection plane, on the other hand, *only* transforms corresponding points above and below the plane into one another. Hence, the 6 operations of a planar triangle *commute* with σ_h.

The symmetry group of the equilateral wedge consists of the original 6 operations of a planar triangle, the horizontal reflection plane, and their products. Since the set with elements $\{E, \sigma_h\}$ forms a group (and each element commutes with the symmetry operations of an equilateral triangle), the appropriate group for the wedge is thereby obtained by taking the direct product

$$\{E, \sigma_{v,1}, \sigma_{v,2}, \sigma_{v,3}, C_3, C_3^2\} \otimes \{E, \sigma_h\} .$$

The 12 elements of this group are

$$\{E, \sigma_{v,1}, \sigma_{v,2}, \sigma_{v,3}, C_3, C_3^2, \sigma_h, \sigma_h\sigma_{v,1}, \sigma_h\sigma_{v,2}, \sigma_h\sigma_{v,3}, \sigma_h C_3, \sigma_h C_3^2\} .$$

∎

6.5.2 Direct Product of Matrices

The determination of the irreducible representations and the character table of a direct product group does not require a separate new calculation of the type discussed in the preceding chapter. Instead, we can utilize the irreducible representations of the two groups used to form the direct to obtain these quantities. To carry out these operations necessitates introducing the direct product of matrices.

The direct product C of two matrices A and B, written as $A \otimes B = C$, is defined in terms of matrix elements by

$$a_{ij} b_{kl} = c_{ik;jl} \,. \tag{6.12}$$

Note that the row and column labels of the matrix elements of C are *composite* labels: the row label, ik, is obtained from the row labels of the matrix elements of A and B and the column label, jl, is obtained from the corresponding column labels. The matrices need not have the same dimension and, in fact, need not even be square. However, since we will apply direct products to construct group representations, we will confine our discussion to square matrices. In this case, if A is an $n \times n$ matrix and B is an $m \times m$ matrix, C is an $mn \times mn$ matrix.

Example 6.2. For matrices A and B given by

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},$$

the direct product $C = A \otimes B$ is

$$A \otimes B = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{12}b_{11} & a_{12}b_{12} & a_{12}b_{13} \\ a_{11}b_{21} & a_{11}b_{22} & a_{11}b_{23} & a_{12}b_{21} & a_{12}b_{22} & a_{12}b_{23} \\ a_{11}b_{31} & a_{11}b_{32} & a_{11}b_{33} & a_{12}b_{31} & a_{12}b_{32} & a_{12}b_{33} \\ a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{22}b_{11} & a_{22}b_{12} & a_{22}b_{13} \\ a_{21}b_{21} & a_{21}b_{22} & a_{21}b_{23} & a_{22}b_{21} & a_{22}b_{22} & a_{22}b_{23} \\ a_{21}b_{31} & a_{21}b_{32} & a_{21}b_{33} & a_{22}b_{31} & a_{22}b_{32} & a_{22}b_{33} \end{pmatrix}.$$

Another way of writing the direct product that more clearly displays its structure is

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix}.$$

∎

The notion of a direct product arises quite naturally in quantum mechanics if we consider the transformation properties of a product of two eigenfunctions. Suppose we have two eigenfunctions φ_i and $\varphi_{i'}$ of a Hamiltonian $\mathcal{H}$ which is invariant under some group of operations. As in Section 6.2, the action of these operations on the eigenfunctions of $\mathcal{H}$ is

$$R\varphi_i = \sum_{j=1}^{\ell} \varphi_j \Gamma_{ji}(R)\,,$$

$$R\varphi_{i'} = \sum_{j'=1}^{\ell'} \varphi_{j'} \Gamma_{j'i'}(R)\,.$$

The question we now ask is: how does the *product* $\varphi_i\varphi_{i'}$ transform under the symmetry operations of the Hamiltonian? Given the transformation properties of φ_i and $\varphi_{i'}$ noted above, we first observe that

$$R(\varphi_i\varphi_{i'}) = R(\varphi_i)R(\varphi_{i'})\,.$$

In other words, since R represents a coordinate transformation, its action on any function of the coordinates is to transform each occurrence of the coordinates. Thus,

$$\begin{aligned} R(\varphi_i\varphi_{i'}) &= \sum_{j=1}^{\ell}\sum_{j'=1}^{\ell'} \varphi_j\varphi_{j'}\Gamma_{ji}(R)\Gamma_{j'i'}(R) \\ &= \sum_{j=1}^{\ell}\sum_{j'=1}^{\ell'} \varphi_j\varphi_{j'}\Gamma_{jj';ii'}(R)\,, \end{aligned}$$

so $\varphi_i\varphi_{i'}$ transforms as the *direct product* of the irreducible representations associated with φ_i and $\varphi_{i'}$.

6.5.3 Representations of Direct Product Groups

Determining the representations of direct products and the construction of their character tables are based on the following theorem:

Theorem 6.1. The direct product of the representations of two groups is a representation of the direct product of these groups.

Proof. A typical product of elements in the direct product group in (6.11) is

$$(a_p b_q)(a_{p'} b_{q'}) = (a_p a_{p'})(b_q b_{q'}) = a_r b_{r'} \,.$$

A representation of the direct product group must preserve the multiplication table. We will use the notation that the matrix $A(a_p b_q)$ corresponds to the element $a_p b_q$. Thus, we must require that

$$A(a_p b_q) A(a_{p'} b_{q'}) = A(a_r b_{r'}) \,.$$

By using the definition of the direct product of two matrices in Equation (6.12), we can write this equation in terms of matrix elements as

$$\begin{aligned}
\Big[A(a_p b_q) A(a_{p'} b_{q'})\Big]_{ik;jl} &= \sum_{m,n} \underbrace{A(a_p b_q)_{ik;mn}}_{A(a_p)_{im} A(b_q)_{kn}} \, \underbrace{A(a_{p'} b_{q'})_{mn;jl}}_{(A(a_{p'})_{mj} A(b_{q'})_{nl}} \\
&= \underbrace{\bigg(\sum_m A(a_p)_{im} A(a_{p'})_{mj}\bigg)}_{A(a_p a_{p'})_{ij}} \underbrace{\bigg(\sum_n A(b_q)]_{kn} A(b_{q'})_{nl}\bigg)}_{A(b_p b_{p'})_{kl}} \\
&= A(a_r)_{ij} A(b_{r'})_{kl} \\
&= A(a_r b_{r'})_{ik;jl} \,.
\end{aligned}$$

Thus, the direct product of the representations preserves the multiplication table of the direct product group and, hence, is a representation of this group. ∎

In fact, as shown in Problem 4 of Problem Set 8, the direct product of irreducible representations of two groups is an irreducible representation of the direct product of those groups. An additional convenient

feature of direct product groups is that the characters of its representations can be computed directly from the characters of the representations of the two groups forming the direct product. This statement is based on the following theorem:

Theorem 6.2. If $\chi(a_p)$ and $\chi(b_q)$ are the characters of representations of two groups G_a and G_b, the characters $\chi(a_p b_q)$ of the representation formed from the matrix direct product of these representations is

$$\chi(a_p b_q) = \chi(a_p)\chi(b_q)\,.$$

Proof. From the definition of the direct product, a representation of the direct product group is

$$A(a_p b_q)_{ij;kl} = A(a_p)_{ik} A(b_q)_{jl}\,.$$

Taking the trace of both sides of this expression yields

$$\underbrace{\sum_{i,j} A(a_p b_q)_{ij;ij}}_{\chi(a_p b_q)} = \underbrace{\Bigl(\sum_i A(a_p)_{ii}\Bigr)}_{\chi(a_p)} \underbrace{\Bigl(\sum_j A(b_q)_{jl}\Bigr)}_{\chi(b_q)}\,.$$

Thus,

$$\chi(a_p b_q) = \chi(a_p)\chi(b_q)\,,$$

which proves the theorem. ∎

Since the characters are associated with a given class, the characters for the classes of the direct product are computed from the characters of the classes of the original groups whose elements contribute to each class of the direct product. Moreover, the number of classes in the direct product group is the product of the numbers of classes in the original groups. This can be seen immediately from the equivalence classes in the direct product group. Using the fact that elements belonging to the different groups commute,

$$(a_i b_j)^{-1}(a_k b_l)(a_i b_j)^{-1} = (a_i^{-1} a_k a_i)(b_j^{-1} b_l b_j)\,.$$

Thus, equivalence classes in the direct product group must be formed from elements in equivalence classes in the original groups.

Example 6.3. Consider the direct product group of the equilateral wedge in Example 6.1. The classes of S_3 are (Example 2.9), in the notation of Example 5.5,

$$\{E\}, \qquad \{\sigma_{v,1}, \sigma_{v,2}, \sigma_{v,3}\}, \qquad \{C_3, C_3^2\}\,,$$

and the classes of the group $\{E, \sigma_h\}$ are

$$\{E\}, \qquad \{\sigma_h\}\,.$$

There are, therefore, six classes in the direct product group, which are obtained by taking the products of elements in the original classes, as discussed above:

$$\begin{gathered}\{E\}, \quad \{C_3, C_3^2\}, \;\; \{\sigma_{v,1}, \sigma_{v,2}, \sigma_{v,3}\}\,, \\ \{\sigma_h\}, \quad \sigma_h\{C_3, C_3^2\}, \;\; \sigma_h\{\sigma_{v,1}, \sigma_{v,2}, \sigma_{v,3}\}\,.\end{gathered}$$

∎

The structure of the character table of the direct product group can now be determined quite easily. We denote the character for the αth class of the jth irreducible representation of group G_a by $\chi_\alpha^j(a_p)$. Similarly, we denote the the character for the βth class of the lth irreducible representation of group G_a by $\chi_\beta^l(b_q)$. Since the direct products of irreducible representations of G_a and G_b are irreducible representations of $G_a \otimes G_b$ (Problem 4, Problem Set 8), and since the classes of $G_a \otimes G_b$ are formed from products of classes of G_a and G_b, the character table of the direct product group has the form

$$\chi_{\alpha\beta}^{jl}(a_p b_q) = \chi_\alpha^j(a_p)\chi_\beta^l(b_q)\,.$$

In other words, with the character tables regarded as square matrices, the character table of the direct product group $G_a \times G_b$ is constructed as a *direct product* of the character tables of G_a and G_b!

Example 6.4. For the direct product group in Example 6.3, the character tables of the original groups are

	E	$3\sigma_v$	$2C_3$
A_1	1	1	1
A_2	1	-1	1
E	2	0	-1

and

	E	σ_h
A_1	1	1
A_2	1	-1

The character table of the direct product group is, therefore, the direct product of these tables (cf. Example 6.1):

	E	$3\sigma_v$	$2C_3$	σ_h	$3\sigma_h\sigma_v$	$2\sigma_h C_3$
A_1^+	1	1	1	1	1	1
A_2^+	1	-1	1	1	-1	1
E^+	2	0	-1	2	0	-1
A_1^-	1	1	1	-1	-1	-1
A_2^-	1	-1	1	-1	1	-1
E^-	2	0	-1	-2	0	1

where the superscript on the irreducible representation refers to the parity under reflection through σ_h. ∎

6.6 Selection Rules

One common application of direct products and their representations is in the determination of selection rules. In this section, we will apply the techniques developed in this chapter to determine the conditions where symmetry requires that a matrix element vanishes.

6.6.1 Matrix Elements

As discussed in Section 6.3.2, the determination of selection rules is based on using group theory to ascertain when the matrix element

$$M_{if} = (i|\mathcal{H}'|f) = \int \varphi_i(x)^* \mathcal{H}' \varphi_f(x) \, \mathrm{d}x \tag{6.13}$$

vanishes by reasons of symmetry. In this matrix element, the initial state transforms according to an irreducible representation $\Gamma^{(i)}$ and the final state transforms according to an irreducible representation $\Gamma^{(f)}$. It only remains to determine the transformation properties of $\mathcal{H}'$. We do this by applying each of the operations of the group of the original Hamiltonian $\mathcal{H}_0$ to the perturbation $\mathcal{H}'$. If we retain only the distinct results of these operations we obtain, by construction, a representation of the group of the Hamiltonian, which we denote by Γ'. This representation may be either reducible or irreducible, depending on $\mathcal{H}'$ and on the symmetry of $\mathcal{H}_0$. If $\mathcal{H}'$ has the same symmetry as $\mathcal{H}_0$, then this procedure generates the identical representation. At the other extreme, if $\mathcal{H}'$ has *none* of the symmetry properties of $\mathcal{H}_0$, then this procedure generates a reducible representation whose dimensionality is equal to the order of the group.

We now consider the symmetry properties under transformation of the product $\mathcal{H}(x)\varphi_f(x)$. From the discussion in Section 6.5.2, we conclude that this quantity transforms as the direct product $\Gamma' \otimes \Gamma^{(f)}$. Since quantities that transform to different irreducible irreducible representations are orthogonal (Problem 6, Problem Set 8), the matrix element (6.13) vanishes if this direct product is either not equal to $\Gamma^{(i)}$ or, if it is reducible, does not include $\Gamma^{(i)}$ in its decomposition. We can summarize this result in the following theorem:

Theorem 6.3. The matrix element

$$(i|\mathcal{H}'|f) = \int \varphi_i(x)^* \mathcal{H}' \varphi_f(x) \, \mathrm{d}x$$

vanishes if the irreducible representation $\Gamma^{(i)}$ corresponding to φ_i is not included in the direct product $\Gamma' \otimes \Gamma^{(f)}$ of the representations Γ' and $\Gamma^{(f)}$ corresponding to $\mathcal{H}'$ and φ_f, respectively.

It is important to note that this selection rule only provides a condition that guarantees that the matrix element will vanish. It does *not* guarantee that the matrix element will not vanish even if the conditions of the theorem are fulfilled.

6.6.2 Dipole Selection Rules

As a scenario which illustrates the power of group theoretical methods, suppose that $\mathcal{H}'$ transforms as a vector, i.e., as (x, y, z). This situation arises when the transitions described by Fermi's Golden Rule (Section 6.3.2) are caused by an electromagnetic field. The form of $\mathcal{H}'$ in the presence of an electromagnetic potential $\boldsymbol{A}$ is obtained by making the replacement[6]

$$\boldsymbol{p} \to \boldsymbol{p} - e\boldsymbol{A}$$

for the momentum in the Hamiltonian. For weak fields, this leads to a perturbation of the form

$$\mathcal{H}' = \frac{e}{m}\boldsymbol{p} \cdot \boldsymbol{A} \tag{6.14}$$

Since the electromagnetic is typically uniform, we can write the matrix element M_{if} as

$$M_{if} \sim (i|\boldsymbol{p}|f) \cdot \boldsymbol{A}$$

so the transformation properties of $\boldsymbol{p} = (p_x, p_y, p_z)$, which are clearly those of a *vector*, determine the selection rules for electromagnetic transitions. These are called the **dipole selection rules.** The examination of many properties of materials rely on the evaluation of dipole matrix elements.

Example 6.6. Suppose the group of the Hamiltonian corresponds to the symmetry operations of an equilateral triangle, i.e., C_{3v}, the character table for which is (Example 5.5)

[6] H. Goldstein, *Classical Mechanics* (Addison–Wesley, Reading, MA, 1950)

C_{3v}	E	$3\sigma_v$	$2C_3$
A_1	1	1	1
A_2	1	-1	1
E	2	0	-1

To determine the dipole selection rules for this system, we must first determine the transformation properties of a vector $\boldsymbol{r} = (x, y, z)$. We take the x- and y-axes in the plane of the equilateral triangle and the z-axis normal to this plane to form a right-handed coordinate system. Applying each symmetry operation to $\boldsymbol{r}$ produces a *reducible* representation because these operations are either rotations or reflections through vertical planes. Thus, the z coordinate is invariant under every symmetry operation of this group which, together with the fact that an (x, y) basis generates the two-dimensional irreducible representation E, yields

$$\Gamma' = A_1 \oplus E$$

We must now calculate the characters associated with the direct products of between Γ' and each irreducible representation to determine the allowed final states given the transformation properties of the initial states. The characters for these direct products are shown below

$\Gamma' = A_1 \oplus E$

C_{3v}	E	$3\sigma_v$	$2C_3$
A_1	1	1	1
A_2	1	-1	1
E	2	0	-1
$A_1 \otimes \Gamma'$	3	1	0
$A_2 \otimes \Gamma'$	3	-1	0
$E \otimes \Gamma'$	6	0	0

Using the decomposition theorem, we find

$$A_1 \otimes \Gamma' = A_1 \oplus E$$

$$A_2 \otimes \Gamma' = A_2 \oplus E$$

$$E \otimes \Gamma' = A_1 \oplus A_2 \oplus 2E$$

Thus, if the initial state transforms as the identical representation A_1, the matrix element vanishes if the final state transforms as A_2. If the initial state transforms as the "parity" representation A_2, the matrix element vanishes if the final state transforms as A_1. Finally, there is no symmetry restriction if the initial state transforms as the "coordinate" representation E. ∎

6.7 Summary

This chapter has demonstrated how the mathematics of groups and their representations are used in quantum mechanics and, indeed, how many of the structures introduced in the preceding chapters appear quite naturally in this context. Apart from exact results, such as Bloch's theorem, we have focussed on the derivation of selection rules induced by perturbations, and derived the principles behind dipole selection rules. A detailed discussion of other applications of discrete groups to quantum mechanical problems is described in the book by Tinkham. Many of the proofs concerning the relation between quantum mechanics and representations of the group of the Hamiltonian are discussed by Wigner.

Chapter 7

Continuous Groups, Lie Groups, and Lie Algebras

Zeno was concerned with three problems . . . These are the problem of the infinitesimal, the infinite, and continuity . . .
—Bertrand Russell

The groups we have considered so far have been, in all but a few cases, discrete and finite. Most of the central theorems for these groups and their representations have relied on carrying out sums over the group elements, often in conjunction with the Rearrangement Theorem (Theorem 2.1). These results provide the basis for the application of groups and representations to physical problems through the construction and manipulation of character tables and the associated computations that require direct sums, direct products, orthogonality and decomposition.

But the notion of symmetry transformations that are based on *continuous* quantities also occur naturally in physical applications. For example, the Hamiltonian of a system with spherical symmetry (e.g., atoms and, in particular, the hydrogen atom) is invariant under all three-dimensional rotations. To address the consequences of this invariance within the framework of group theory necessitates confronting several issues that arise from the continuum of rotation angles. These include defining what we mean by a "multiplication table," determining how summations over group elements are carried out, and deriving

the appropriate re-statement of the Rearrangement Theorem to enable the Great Orthogonality Theorem and its consequences to be obtained for continuous groups. More generally, the existence of a continuum of group elements, when combined with the requirement of analyticity, introduces new structures associated with constructing differentials and integrals of group elements. In effect, this represents an amalgamation of group theory and analysis, so such groups are the natural objects for describing the symmetry of analytic structures such as differential equations and those that arise in differential geometry. In fact, the introduction of analytic groups by Sophus Lie late in the 19th century was motivated by the search for symmetries of differential equations.

In this chapter we begin our discussion about the modifications to our development of groups and representations that are necessitated by having a continuum of elements. We begin in the first section with the definition of a continuous group and specialize to the most common type of continuous group, the Lie group. We then introduce the idea of an infinitesimal generator of a transformation, from which every element can be obtained by repeated application. These generators embody much of the structure of the group and, because there are a finite number of these entities, are simpler to work with than the full group. This leads naturally to the Lie algebra associated with a Lie group. All of these concepts are illustrated with the groups of proper rotations in two and three dimensions. The representation of these groups, their character tables, and basis functions will be discussed in the next chapter.

7.1 Continuous Groups

Consider a set of elements R that depend on a number of real continuous parameters, $R(a) \equiv R(a_1, a_2, \ldots, a_r)$. These elements are said to form a **continuous group** if they fulfill the requirements of a group (Section 2.1) and if there is some notion of 'proximity' or 'continuity' imposed on the elements of the group in the sense that a small change in one of the factors of a product produces a correspondingly small change in their product. If the group elements depend on r parameters, this is called an r-parameter continuous group.

In general terms, the requirements that a continuous set of elements form a group are the same as those for discrete elements, namely, closure under multiplication, associativity, the existence of a unit, and an inverse for every element. Consider first the multiplication of two elements $R(a)$ and $R(b)$ to yield the product $R(c)$:

$$R(c) = R(a)R(b)\,.$$

Then c must be a continuous real function f of a and b:

$$c = f(a,b)\,.$$

This defines the structure of the group in the same way as the multiplication table does for discrete groups. The associativity of the composition law,

$$R(a)\underbrace{\Big[R(b)R(c)\Big]}_{R[f(b,c)]} = \underbrace{\Big[R(a)R(b)\Big]}_{R[f(b,c)]}R(c)\,,$$

requires that

$$f[a,f(b,c)] = f[f(a,b),c]\,.$$

The existence of an identity element, which we denote by $R(a_0)$,

$$R(a_0)R(a) = R(a)R(a_0) = R(a)\,,$$

is expressed in terms of f as

$$f(a_0,a) = f(a,a_0) = a\,.$$

The inverse of each element $R(a)$, denoted by $R(a')$, produces

$$R(a')R(a) = R(a)R(a') = R(a_0)\,.$$

Therefore,

$$f(a',a) = f(a,a') = a_0\,.$$

If f is an analytic function, i.e., a function with a convergent Taylor series expansion within the domain defined by the parameters, the

resulting group is called an r-parameter **Lie group**, named after Sophus Lie, a Norwegian mathematician who provided the foundations for such groups.

Our interest in physical applications centers around transformations on d-dimensional spaces. Examples include Euclidean spaces, where the variables are spatial coordinates, Minkowski spaces, where the variables are space-time coordinates, and spaces associated with internal degrees of freedom, such as spin or isospin. In all cases, these are mappings of the space onto itself and have the general form

$$x_i' = f_i(x_1, x_2, \ldots, x_d; a_1, a_2, \ldots, a_r), \qquad i = 1, 2, \ldots, d\,.$$

If the f_i are analytic, then this defines an r-parameter Lie group of transformations.

Example 7.1 Consider the one-dimensional transformations

$$x' = ax \tag{7.1}$$

where a is an non-zero real number. This transformation corresponds to stretching the real line by a factor a. The product of two such operations, $x'' = ax'$ and $x' = bx$ is

$$x'' = ax' = abx\,.$$

By writing $x'' = cx$, we have that

$$c = ab\,, \tag{7.2}$$

so the multiplication of two transformations is described by an analytic function that yields another transformation of the form in (7.1). This operation is clearly associative, as well as Abelian, since the product transformation corresponds to the multiplication of real numbers. This product can also be used to determine the inverse of these transformations. By setting $c = 1$ in (7.2), so that $x'' = x$, the inverse of (7.1) is seen to correspond to the transformation with $a' = a^{-1}$, which explains the requirement that $a \neq 0$. Finally, the identity is determined from $x' = x$, which clearly corresponds to the transformation

with $a = 1$. Hence, the transformations defined in (7.1) form a one-parameter Abelian Lie group. ▮

Example 7.2 Now consider the one-dimensional transformations

$$x' = a_1 x + a_2\,, \tag{7.3}$$

where again a_1 is an non-zero real number. These transformations corresponds to the stretching of the real line by a factor a_1, as in the preceding Example, and a translation by a_2. The product of two operations is

$$x'' = a_1 x' + a_2 = a_1(b_1 x + b_2) + a_2 = a_1 b_1 x + a_1 b_2 + a_2\,.$$

By writing $x'' = c_1 x + c_2$, we have that

$$c_1 = a_1 b_1, \qquad c_2 = a_1 b_2 + a_2\,,$$

so the multiplication of two transformations is described by an analytic function and yields another transformation of the form in (7.1). However, although this multiplication is associative, it is not Abelian, as can be seen from the fact that the indices do not enter symmetrically in c_2. By setting, $c_1 = c_2 = 1$, the inverse of (7.3) is the transformation

$$x' = \frac{x}{a_1} - \frac{a_2}{a_1}\,.$$

The identity is again determined from $x' = x$, which requires that $a_1 = 1$ and $a_2 = 0$. Hence, the transformations in (7.3) form a two-parameter (non-Abelian) Lie group. ▮

7.2 Linear Transformation Groups

An important class of transformations is the group of linear transformations in d dimensions. These can be represented by $d \times d$ matrices. For example, the most general such transformation in two dimensions is $\boldsymbol{x}' = A\boldsymbol{x}$ or, in matrix form,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\,, \tag{7.4}$$

where $\det(A) = a_{11}a_{22} - a_{12}a_{21} \neq 0$ (Example 2.4). With no further restriction, and with the composition of two elements given by the usual rules of matrix multiplication, these matrices form a four-parameter Lie group. This Lie group is called the **general linear group** in two dimensions and is denoted by GL(2,R), where the 'R' signifies that the entries are real; the corresponding group with complex entries is denoted by GL(2,C). In n dimensions, these transformation groups are denoted by GL(n,R), or, with complex entries, by GL(n,C).

7.2.1 Orthogonal Groups

Many transformations in physical applications are required to preserve length in the appropriate space. If that space is ordinary Euclidean n-dimensional space, the restriction that lengths be preserved means that

$$x_1'^2 + x_2'^2 + \cdots + x_n'^2 = x_1^2 + x_2^2 + \cdots + x_n^2 \,. \tag{7.5}$$

The corresponding groups, which are subgroups of the general linear group, are called **orthogonal**, and are denoted by O(n).

Consider the orthogonal group in two-dimensions, i.e., O(2), where the coordinates are x and y. By substituting the general transformation (7.4) into (7.5), we require that

$$\begin{aligned} x'^2 + y'^2 &= (a_{11}x + a_{12}y)^2 + (a_{21}x + a_{22}y)^2 \\ &= (a_{11}^2 + a_{21}^2)x^2 + 2(a_{11}a_{12} + a_{21}a_{22})xy + (a_{12}^2 + a_{22}^2)y^2 \,. \end{aligned}$$

For the right-hand side of this equation to be equal to $x^2 + y^2$ for *all* x and y, we must set

$$a_{11}^2 + a_{21}^2 = 1, \qquad a_{11}a_{12} + a_{21}a_{22} = 0, \qquad a_{12}^2 + a_{22}^2 = 1 \,.$$

Thus, we have three conditions imposed on four parameters, leaving one free parameter. These conditions can be used to establish the following relation:

$$(a_{11}a_{22} - a_{12}a_{21})^2 = 1 \,.$$

Recognizing the quantity in parentheses as the determinant of the transformation, this condition implies that

$$\det(A) = \pm 1\,.$$

If $\det(A) = 1$, then the parity of the coordinate system is not changed by the transformation; this corresponds to a *proper* rotation. If $\det(A) = -1$, then the parity of the coordinate system is changed by the transformation; this corresponds to an *improper* rotation. As we have already seen, both types of transformations are important in physical applications, but we will first examine the proper rotations in two-dimensions. This group is called the *special* orthogonal group in two dimensions and is denoted by SO(2), where "special" signifies the restriction to proper rotations. The parametrization of this group that we will use is

$$R(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \tag{7.6}$$

where φ, the single parameter in this Lie group, is the rotation angle of the transformation. As can easily be checked using the trigonometric identities for the sum of two angles,

$$R(\varphi_1 + \varphi_2) = R(\varphi_1)R(\varphi_2)\,, \tag{7.7}$$

so this group is clearly Abelian.

7.3 Infinitesimal Generators

A construction of immense utility in the study of Lie groups, which was introduced and extensively studied by Lie, is the **infinitesimal generator**. The idea behind this is that instead of having to consider the group as a whole, for many purposes it is sufficient to consider an infinitesimal transformation around the identity. Any finite transformation can then be constructed by the repeated application, or "integration," of this infinitesimal transformation.

7.3.1 Matrix Form of Generators

For SO(2), we first expand $R(\varphi)$ in a Taylor series around the identity ($\varphi = 0$):

$$R(\varphi) = R(0) + \frac{\mathrm{d}R}{\mathrm{d}\varphi}\bigg|_{\varphi=0}\varphi + \frac{1}{2}\frac{\mathrm{d}^2R}{\mathrm{d}\varphi^2}\bigg|_{\varphi=0}\varphi^2 + \cdots\,. \tag{7.8}$$

The coefficients in this series can be determined directly from (7.6), but a more elegant solution may be found by first differentiating (7.7) with respect to φ_1,

$$\frac{\mathrm{d}}{\mathrm{d}\varphi_1}R(\varphi_1 + \varphi_2) = \frac{\mathrm{d}R(\varphi_1)}{\mathrm{d}\varphi_1}R(\varphi_2)\,, \tag{7.9}$$

then setting $\varphi_1 = 0$. Using the chain rule, the left-hand side of this equation is

$$\left[\frac{\mathrm{d}R(\varphi_1 + \varphi_2)}{\mathrm{d}(\varphi_1 + \varphi_2)}\,\frac{\mathrm{d}(\varphi_1 + \varphi_2)}{\mathrm{d}\varphi_1}\right]\bigg|_{\varphi_1=0} = \frac{\mathrm{d}R(\varphi_2)}{\mathrm{d}\varphi_2}\,,$$

so Eq. (7.9) becomes

$$\frac{\mathrm{d}R(\varphi)}{\mathrm{d}\varphi} = XR(\varphi)\,, \tag{7.10}$$

where

$$\frac{\mathrm{d}R(\varphi_1)}{\mathrm{d}\varphi_1}\bigg|_{\varphi_1=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \equiv X\,. \tag{7.11}$$

Equations (7.10) and (7.11) allow us to determine all of the expansions coefficients in (7.9). By setting $\varphi = 0$ in (7.10) and observing that $R(0) = I$, where I is the 2×2 unit matrix,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\,,$$

we obtain

$$\frac{\mathrm{d}R(\varphi)}{\mathrm{d}\varphi}\bigg|_{\varphi=0} = X\,. \tag{7.12}$$

To determine the higher-order derivatives of R, we differentiate (7.10) n times, and set $\varphi = 0$:

$$\frac{\mathrm{d}^n R(\varphi)}{\mathrm{d}\varphi^n}\bigg|_{\varphi=0} = X \frac{\mathrm{d}^{n-1} R(\varphi)}{\mathrm{d}\varphi^{n-1}}\bigg|_{\varphi=0} .$$

This yields, in conjunction with (7.12),

$$\frac{\mathrm{d}^n R(\varphi)}{\mathrm{d}\varphi^n}\bigg|_{\varphi=0} = X^n .$$

Substituting this expression into the Taylor series in (7.8) allows us to write

$$\begin{aligned} R(\varphi) &= I + X\varphi + \tfrac{1}{2}X^2\varphi^2 + \cdots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}(X\varphi)^n \\ &\equiv \mathrm{e}^{\varphi X} , \end{aligned}$$

where $X^0 = I$ and the exponential of a matrix is *defined* by the Taylor series expansion of the exponential. Thus, every rotation by a finite angle can be obtain from the exponentiation of the matrix X, which is called the **infinitesimal generator** of rotations. Since $X^2 = I$, it is a straightforward matter to show directly from the Taylor series of the exponential (Problem 4, Problem Set 9) that

$$\mathrm{e}^{\varphi X} = I\cos\varphi + X\sin\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} .$$

7.3.2 Operator Form of Generators

An alternative way of representing infinitesimal generators through which connections with quantum mechanics can be directly made is in terms of differential operators. To derive the operator associated with infinitesimal rotations, we expand (7.6) to first order in $\mathrm{d}\varphi$ to obtain the transformation

$$\begin{aligned} x' &= x\cos\varphi - y\sin\varphi = x - y\,\mathrm{d}\varphi , \\ y' &= x\sin\varphi + y\cos\varphi = x\,\mathrm{d}\varphi + y . \end{aligned}$$

An arbitrary differentiable function $F(x,y)$ then transforms as

$$F(x',y') = F(x - y\,\mathrm{d}\varphi, x\,\mathrm{d}\varphi + y)\,.$$

Retaining terms to first order in $\mathrm{d}\varphi$ on the right-hand side of this equation yields

$$F(x',y') = F(x,y) + \Big(-y\frac{\partial F}{\partial x} + x\frac{\partial F}{\partial y}\Big)\,\mathrm{d}\varphi\,.$$

Since F is an arbitrary function, we can associate infinitesimal rotations with the operator

$$X = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\,.$$

As we will see in the next section, this operator is proportional to the z-component of the angular momentum operator.

The group SO(2) is simple enough that the full benefits of an infinitesimal generator are not readily apparent. We will see in the next section, where we discuss SO(3), that the infinitesimal generators embody much of the structure of the full group.

7.4 SO(3)

The orthogonal group in three dimensions is comprised of the transformations that leave the quantity $x^2 + y^2 + z^2$ invariant. The group GL(3,R) has 9 parameters, but the invariance of the length produces six independent conditions, leaving three free parameters, so O(3) forms a three-parameter Lie group. If we restrict ourselves to transformations with unit determinant, we obtain the group of proper rotations in three dimensions, SO(3).

There are three common ways to parametrize these rotations:

- Successive rotations about three mutually orthogonal *fixed* axes.
- Successive about the z-axis, about the *new* y-axis, and then about the *new* z-axis. These are called **Euler angles**.

- The axis-angle representation, defined in terms of an axis whose direction is specified by a unit vector (two parameters) and a rotation about that axis (one parameter).

In this section, we will use the first of these parametrizations to demonstrate some of the properties of SO(3). In the next chapter, where we will develop the orthogonality relations for this group, the axis-angle representation will prove more convenient.

7.4.1 Rotation Matrices

Consider first rotations about the z-axis by an angle φ_3:

$$R_3(\varphi_3) = \begin{pmatrix} \cos\varphi_3 & -\sin\varphi_3 & 0 \\ \sin\varphi_3 & \cos\varphi_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

The corresponding infinitesimal generator is calculated as in (7.11):

$$X_3 = \frac{\mathrm{d}R_3}{\mathrm{d}\varphi_3}\bigg|_{\varphi_3=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

These results are essentially identical to those for SO(2). However, for SO(3), we have rotations about two other axes to consider. For rotations about the x-axes by an angle φ_1, the rotation matrix is

$$R_1(\varphi_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_1 & -\sin\varphi_1 \\ 0 & \sin\varphi_1 & \cos\varphi_1 \end{pmatrix}$$

and the corresponding generator is

$$X_1 = \frac{\mathrm{d}R_1}{\mathrm{d}\varphi_1}\bigg|_{\varphi_1=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

Finally, for rotations about the y-axis by an angle φ_2, we have

$$R_2(\varphi_2) = \begin{pmatrix} \cos\varphi_2 & 0 & \sin\varphi_2 \\ 0 & 1 & 0 \\ -\sin\varphi_2 & 0 & \cos\varphi_2 \end{pmatrix}$$

and the generator is

$$X_2 = \frac{\mathrm{d}R_2}{\mathrm{d}\varphi_2}\bigg|_{\varphi_2=0} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

As can be easily verified, the matrices $R_i(\varphi_i)$ do not commute, nor do the X_i. However, the X_i have an additional useful property, namely closure under commutation. As an example, consider the products X_1X_2 and X_2X_1:

$$X_1X_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$X_2X_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus, the commutator of X_1 and X_2, denoted by $[X_1, X_2]$ is given by

$$[X_1, X_2] \equiv X_1X_2 - X_2X_1 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = X_3$$

Similarly, we have

$$[X_2, X_3] = X_1, \qquad [X_3, X_1] = X_2$$

The commutation relations among all of the X_i can be succinctly summarized by introducing the anti-symmetric symbol ε_{ijk}, which takes the

value 1 for a symmetric permutation of distinct i, j, and k, the value -1 for an antisymmetric permutation, and is zero otherwise (i.e., if two or more of i, j and k are equal). We can then write

$$[X_i, X_j] = \varepsilon_{ijk} X_k \tag{7.13}$$

We will discuss the physical interpretation of these generators once we obtain their operator form in the next section.

7.4.2 Operators for Infinitesimal Rotations

As was the case in Section 7.3, an alternative to the matrix representation of infinitesimal generators is in terms of differential operators. Proceeding as in that section, we first write the general rotation as an expansion to first order in each of the φ_i about the identity. This yields the transformation matrix

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & -\varphi_3 & \varphi_2 \\ \varphi_3 & 1 & -\varphi_1 \\ -\varphi_2 & \varphi_1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Substituting this coordinate transformation into a differentiable function $F(x, y, z)$,

$$F(x', y', z') = F(x - \varphi_3 y + \varphi_2 z, y + \varphi_3 x - \varphi_1 z, z - \varphi_2 x + \varphi_1 y)$$

and expanding the right-hand side to first order in the φ_i yields the following expression:

$$\begin{aligned} F(x', y', z') &= F(x, y, z) \\ &+ \Big(\frac{\partial F}{\partial z} y - \frac{\partial F}{\partial y} z\Big)\varphi_1 + \Big(\frac{\partial F}{\partial x} z - \frac{\partial F}{\partial z} x\Big)\varphi_2 + \Big(\frac{\partial F}{\partial y} x - \frac{\partial F}{\partial x} y\Big)\varphi_3 \end{aligned}$$

Since F is an arbitrary differentiable function, we can identify the generators X_i of rotations about the coordinate axes from the coefficients of the φ_i, i.e., with the differential operators

$$X_1 = y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}$$

$$X_2 = z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \tag{7.14}$$

$$X_3 = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$$

Notice that X_3 is the operator obtained for SO(2) in Section 7.3. We can now assign a physical interpretation to these operators by comparing them with the vectors components of the angular operators in the coordinate representation, obtained from the definition

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p} = \boldsymbol{r} \times (-i\hbar\boldsymbol{\nabla})$$

Carrying out the cross-product yields the standard expressions

$$L_1 = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$

$$L_2 = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \tag{7.15}$$

$$L_3 = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

for the x, y, and z components of $\boldsymbol{L}$, respectively. Thus, $L_i = -i\hbar X_i$, for $i = 1, 2, 3$, and (7.13) becomes

$$[L_i, L_j] = i\hbar\varepsilon_{ijk}L_k$$

which are the usual angular momentum commutation relations. Therefore, we can associate the vector components of the angular momentum operator with the generators of infinitesimal rotations about the corresponding axes. An analogous association exists between the vector components of the coordinate representation of the linear momentum operator and differential translation operations along the corresponding directions.

7.4.3 The Algebra of Infinitesimal Generators

The commutation relations in (7.13) define a "product" of two generators which yields the third generator. Thus, the set of generators is

closed under this operation. Triple products, which determine whether or not this composition law is associative, can be written in a concise form using only the definition of the commutator, i.e., in the form of an identity, without any explicit reference to the quantities involved. Beginning with the triple product

$$\begin{aligned}\big[A,[B,C]\big] &= A[B,C]-[B,C]A\\ &= ABC-ACB-BCA+CBA\end{aligned}$$

We now add and subtract the quantities BAC and CAB on the right-hand side of this equation and rearrange the resulting expression into commutators to obtain

$$\begin{aligned}\big[A,[B,C]\big] &= ABC-ACB-BCA+CBA\\ &\quad +BAC-BAC+CAB-CAB\\ &= -C(AB-BA)+(AB-BA)C\\ &\quad +B(AC-CA)-(AC-CA)B\\ &= -\big[[A,B],C\big]+\big[[C,A],B\big]\end{aligned}$$

A simple rearrangement yields the **Jacobi identity**:

$$\big[A,[B,C]\big]+\big[B,[C,A]\big]+\big[C,[A,B]\big]=0$$

Notice that this identity has been obtained using only the definition of the commutator.

For the infinitesimal generators of the rotation group, with the commutator in (7.13), each of the terms in the Jacobi identity vanishes. Thus,

$$\big[A,[B,C]\big]=\big[[A,B],C\big]$$

so the product of these generators is associative. In the more general case, however, products of quantities defined in terms of a commutator are not associative. The **Lie algebra** associated with the Lie group from

which the generators are obtained consists of quantities $A, B, C, \ldots$ defined by

$$A = \sum_{k=1}^{3} a_k X_k, \qquad B = \sum_{k=1}^{3} b_k X_k, \qquad C = \sum_{k=1}^{3} c_k X_k, \qquad \text{etc.}$$

where the $a_k, b_k, c_k, \ldots$ are real coefficients and from which linear combinations $\alpha A + \beta B$ with real α and β can be formed. The product is given by

$$[A, B] = -[B, A]$$

and the Jacobi identity is, of course, satisfied.

The formal definition of a Lie algebra, which is an abstraction of the properties just discussed, is as follows.

Definition. A **Lie algebra** is a vector space L over some field F[1] (typically the real or complex numbers) together with a binary operation $[\cdot, \cdot] : L \times L \to L$, called the *Lie bracket*, which has the following properties:

1. **Bilinearity**.

$$\begin{aligned}[ax + by, z] &= a[x, z] + b[y, z] \\ [z, ax + by] &= a[z, x] + b[z, y]\end{aligned}$$

   for all a and b in F and x, y, and z in L.

2. **Jacobi identity**.

$$\big[[x, y], z\big] + \big[[z, x], y\big] + \big[[y, z], x\big] = 0$$

   for all x, y, and z in L.

[1]A field is an algebraic system of elements in which the operations of addition, subtraction, multiplication, and division (except by zero) may be performed without leaving the system (closure) and the associative, commutative, and distributive rules, familiar from the arithmetic of ordinary numbers, hold. Examples of fields are the rational numbers, the real numbers, and the complex numbers. The smallest field has only two elements: $\{0, 1\}$. The concept of a field is useful for defining vectors and matrices, whose components can be elements of any field.

3. **Antisymmetry**.

$$[x, y] = -[y, x]$$

for all x and y in L.

7.5 Summary

In this chapter, we have described the properties of Lie groups in terms of specific examples, especially SO(2) and SO(3). With this background, we can generalize our discussion to any Lie group. An r-parameter Lie group of transformations on an n-dimensional space is

$$x_i' = f_i(x_1, x_2, \dots, x_n; a_1, a_2, \dots, a_r)$$

where $i = 1, 2, \dots, n$. If only one of the r parameters a_i is changed from zero, while all the other parameters are held fixed, we obtain the infinitesimal transformations X_i associated with this Lie group. These can be expressed as differential operators by examining the effect of these infinitesimal coordinate transformations on an arbitrary differentiable function F:

$$\begin{aligned} \mathrm{d}F &= \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}\,\mathrm{d}x_j \\ &= \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}\biggl(\sum_{i=1}^{r} \frac{\partial f_j}{\partial a_i}\bigg|_{a=0} \mathrm{d}a_i\biggr) \\ &= \sum_{i=1}^{r} \mathrm{d}a_i\biggl(\sum_{j=1}^{n} \frac{\partial f_j}{\partial a_i}\bigg|_{a=0} \frac{\partial}{\partial x_j}\biggr) F \end{aligned}$$

We identify the differential operators X_i as the coefficient of $\mathrm{d}a_i$ in this differential:

$$X_i = \sum_{j=1}^{n} \frac{\partial f_j}{\partial a_i}\bigg|_{a=0} \frac{\partial}{\partial x_j}$$

for $r = 1, 2, \dots, r$. These operators satisfy commutation relations of the form

$$[X_i, X_j] = c_{ij}^k X_k$$

where the c_{ij}^k are called **structure constants** and are a property of the group. The commutator satisfies the Jacobi identity,

$$\Big[X_i, [X_j, X_k]\Big] + \Big[X_j, [X_k, X_i]\Big] + \Big[X_k, [X_i, X_j]\Big] = 0$$

which places a constraint on the structure constants. The commutator and the Jacobi identity, together with the ability to form real linear combinations of the X_i endows these generators with the structure of an algebra, called the Lie algebra associated with the Lie group.

Chapter 8

Irreducible Representations of SO(2) and SO(3)

The shortest path between two truths in the real domain passes through the complex domain.

—Jacques Hadamard[1]

Some of the most useful aspects of group theory for applications to physical problems stem from the orthogonality relations of characters of irreducible representations. The widespread impact of these relations stems from their role in constructing and resolving new representations from direct products of irreducible representations. Direct products are especially important in applications involving continuous groups, with the construction of higher dimensional irreducible representations, the derivation of angular momentum coupling rules, and the characterization of families of elementary particles all relying on the formation and decomposition of direct products.

Although the notion of an irreducible representation can be carried over directly from our development of discrete groups through Schur's first lemma, a transcription of Schur's second lemma and the Great Orthogonality Theorem to the language of continuous groups requires a separate discussion. This is because proving the latter two theorems

[1]Quoted in *The Mathematical Intelligencer* **13**(1), 1991.

necessitates performing summations over group elements and invoking the Rearrangement Theorem (Theorem 2.1). This theorem guarantees the following equality

$$\sum_g f(g) = \sum_g f(g'g)\,, \tag{8.1}$$

where the summation is over elements g in a group G, g' is any other element in G, and f is some function of the group elements. The crucial point is that the same quantities appear on both sides of the equation; the only difference is the order of their appearance. To proceed with the proofs of these theorems for continuous groups requires an equality analogous to (8.1):

$$\int f(R)\,\mathrm{d}R = \int f(R'R)\,\mathrm{d}R\,, \tag{8.2}$$

where R and R' are the elements of a continuous group and f is some function of these elements. To appreciate the issues involved, we write the integral on the left-hand side of (8.2) as an integral over the parameters

$$\int f(R)\,\mathrm{d}R = \int f(R)g(R)\,\mathrm{d}a\,, \tag{8.3}$$

where $g(R)$ is the density of group elements in parameter space in the neighborhood of R. The equality in (8.2) will hold provided that the density of group elements is arranged so that the density of the points $R'R$ is the same as that of the points R. Our task is to find the form of $g(R)$ which ensures this. A related concept that will arise is the notion of the "order" of the continuous group as the volume of its elements in the space defined by the parameters of the group.

This chapter is devoted to the characters and irreducible representations of SO(2) and SO(3). For SO(2), we will show that the density of group elements is uniform across parameter space, so the density function reduces to a constant. But, for SO(3), we will need to carry out the determination of the density function in (8.3) explicitly. This will illustrate the general procedure which is applicable to any group. For both SO(2) and SO(3), we will derive the basis functions for their irreducible representations which will be used to obtain the corresponding characters and to demonstrate their orthogonality

8.1 Orthogonality of Characters for SO(2)

The structure of SO(2) is simple enough that many of the results obtained for discrete groups can be taken over directly with little or no modification. The basis of this claim is that the Rearrangement Theorem for this group is, apart from the replacement of the sum by an integral, a direct transcription of that for discrete groups which, together with this group being Abelian, renders the calculation of characters a straightforward exercise.

8.1.1 The Rearrangement Theorem

We first show that the rearrangement theorem for this group is

$$\int_0^{2\pi} R(\varphi')R(\varphi)\,\mathrm{d}\varphi = \int_0^{2\pi} R(\varphi)\,\mathrm{d}\varphi\,.$$

This implies that the weight function appearing in (8.3) is unity, i.e., the density of group elements is uniform in the space of the parameter φ. Using the fact that $R(\varphi')R(\varphi) = R(\varphi' + \varphi)$, we have

$$\int_0^{2\pi} R(\varphi')R(\varphi)\,\mathrm{d}\varphi = \int_0^{2\pi} R(\varphi' + \varphi)\,\mathrm{d}\varphi\,.$$

We now introduce a new integration variable $\theta = \varphi' + \varphi$. Since φ' is fixed, we have that $\mathrm{d}\varphi = \mathrm{d}\theta$. Then, making the appropriate changes in the upper and lower limits of integration, and using the fact that $R(\varphi + 2\pi) = R(\varphi)$, yields

$$\begin{aligned}
\int_0^{2\pi} R(\varphi' + \varphi)\,\mathrm{d}\varphi &= \int_{\varphi'}^{\varphi'+2\pi} R(\theta)\,\mathrm{d}\theta \\
&= \int_{\varphi'}^{2\pi} R(\theta)\,\mathrm{d}\theta + \int_{2\pi}^{\varphi'+2\pi} R(\theta)\,\mathrm{d}\theta \\
&= \int_{\varphi'}^{2\pi} R(\theta)\,\mathrm{d}\theta + \int_0^{\varphi'} R(\theta)\,\mathrm{d}\theta \\
&= \int_0^{2\pi} R(\theta)\,\mathrm{d}\theta\,,
\end{aligned}$$

which verifies our assertion.

8.1.2 Characters of Irreducible Representations

We can now use Schur's first lemma for SO(2). Since SO(2) is an Abelian group, this first lemma requires all of the irreducible representations to be one-dimensional (cf. Problem 4, Problem Set 5). Thus, every element is in a class by itself and the characters must satisfy the same multiplication rules as the elements of the group:

$$\chi(\varphi)\chi(\varphi') = \chi(\varphi + \varphi') . \tag{8.4}$$

The character corresponding to the unit element, $\chi(0)$, which must map onto the identity for ordinary multiplication, is clearly unity for all irreducible representations:

$$\chi(0) = 1 . \tag{8.5}$$

Finally, we require the irreducible representations to be single-valued, i.e., an increase in the rotation angle by 2π does not change the effect of the rotation. Thus,

$$\chi(\varphi + 2\pi) = \chi(\varphi) . \tag{8.6}$$

The three conditions in (8.4), (8.5), and (8.6) are sufficient to determine the characters of all of the irreducible representations of SO(2).

We will proceed by writing Eq. (8.4) as a differential equation and using (8.5) as an "initial condition" and (8.6) as a "boundary condition." In (8.4), we set $\varphi' = \mathrm{d}\varphi$,

$$\chi(\varphi)\chi(\mathrm{d}\varphi) = \chi(\varphi + \mathrm{d}\varphi) ,$$

and expand both sides of this equation to first order in $\mathrm{d}\varphi$:

$$\chi(\varphi)\left[\chi(0) + \frac{\mathrm{d}\chi}{\mathrm{d}\varphi}\bigg|_{\varphi=0}\mathrm{d}\varphi\right] = \chi(\varphi) + \frac{\mathrm{d}\chi}{\mathrm{d}\varphi}\mathrm{d}\varphi .$$

Then, using (8.5) and cancelling common terms, this equation reduces to a first-order ordinary differential equation for $\chi(\varphi)$:

$$\frac{\mathrm{d}\chi}{\mathrm{d}\varphi} = \chi_0'\chi(\varphi) ,$$

where $\chi_0' = \chi'(0)$ is to be determined. The general solution to this equation is

$$\chi(\varphi) = A\,\mathrm{e}^{\chi_0'\varphi}\,,$$

where A is a constant of integration which is also to be determined. In fact, by setting $\varphi = 0$ and invoking (8.5), we see that $A = 1$. The requirement (8.6) of single-valuedness, when applied to this solution, yields the condition that

$$\mathrm{e}^{\chi_0'(\varphi+2\pi)} = \mathrm{e}^{\chi_0'\varphi}\,,$$

or,

$$\mathrm{e}^{2\pi\chi_0'} = 1\,.$$

The most general solution of this equation is $\chi_0' = im$, where $i^2 = -1$ and m is any integer. This produces an infinite sequence of characters of the irreducible representations of SO(2):

$$\chi^{(m)}(\varphi) = \mathrm{e}^{im\varphi}, \qquad m = \ldots, -2, -1, 0, 1, 2, \ldots \tag{8.7}$$

The identical representation corresponds to $m = 0$. In contrast to the case of finite groups, we see that SO(2) has an infinite set of irreducible representations, albeit one that is countably infinite.

8.1.3 Orthogonality Relations

Having determined the characters for SO(2), we can now examine the validity of the orthogonality theorems for characters which were discussed for discrete groups in Theorem 5.1. We proceed heuristically and begin by observing that the exponential functions in (8.7) are orthogonal over the interval $0 \le \varphi < 2\pi$:

$$\int_0^{2\pi} \mathrm{e}^{i(m'-m)\varphi}\,\mathrm{d}\varphi = 2\pi\delta_{m,m'}\,.$$

By writing this relation as

$$\int_0^{2\pi} \chi^{(m)*}(\varphi)\chi^{(m')}(\varphi)\,\mathrm{d}\varphi = 2\pi\delta_{m,m'}\,, \tag{8.8}$$

we obtain an orthogonality relation of the form in Eq. (5.4), once we identify the "order" of SO(2) as the quantity

$$\int_0^{2\pi} d\varphi = 2\pi\,.$$

This is the "volume" of the group in the space of the parameter φ, which lies in the range $0 \le \varphi < 2\pi$, given that the density function is unity, according to the discussion in thew preceding section. Note that the integration over φ is effectively a sum over classes.

Example 8.1. Consider the representation of SO(2) derived in Section 7.2:

$$R(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\,. \tag{8.9}$$

Since SO(2) is an Abelian group, this representation must be reducible. We can decompose this representation into its irreducible components by using either the analogue of the Decomposition Theorem (Section 5.3) for continuous groups or, more directly, by using identities between complex exponential and trigonometric functions:

$$\begin{aligned} \chi(\varphi) &\equiv \mathrm{tr}[R(\varphi)] \\ &= 2\cos\varphi \\ &= e^{i\varphi} + e^{-i\varphi}\,. \end{aligned}$$

A comparison with (8.7) yields

$$\chi(\varphi) = \chi^{(1)}(\varphi) + \chi^{(-1)}(\varphi)\,,$$

so the representation in (8.9) is a direct sum of the irreducible representations corresponding to $m = 1$ and $m = -1$.[2] ▮

[2]This example illustrates the importance of the field used in the entries of the matrices for SO(2). If we are restricted to *real* entries, then the representation in (8.9) is *irreducible*. But, if the entries are *complex*, then this example shows that this representation is *reducible*.

8.2 Basis Functions for Irreducible Representations

We were able to determine the characters for all of the irreducible representations of SO(2) without any knowledge of the representations themselves. But this is not the typical case for continuous groups. We will see, for example, when determining the characters for SO(3) that we will be required to construct explicit representations of rotations corresponding to different classes. The action of these rotations on the basis functions will determine the representation of that class and the character will be calculated directly from this representation. As an introduction to that discussion, in this section we will determine the basis functions of the irreducible representations of SO(2).

We begin by calculating the eigenvalues of the matrix in (8.9) from $\det(R - \lambda I) = 0$:

$$\begin{aligned} \begin{vmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{vmatrix} &= (\cos\varphi - \lambda)^2 + \sin^2\varphi \\ &= \lambda^2 - 2\lambda\cos\varphi + 1 = 0\,. \end{aligned}$$

Solving for λ yields

$$\lambda = \cos\varphi \pm i\sin\varphi = \mathrm{e}^{\pm i\varphi}\,. \tag{8.10}$$

The corresponding eigenvectors are proportional to $x \pm iy$. Thus, operating on these eigenvectors with $R(\varphi)$ (see below) generates the irreducible representations corresponding to $m = 1$ and $m = -1$ in (8.7), i.e., the characters $\chi^{(1)}(\varphi)$ and $\chi^{(-1)}(\varphi)$.

Obtaining the basis functions for the other irreducible representations of SO(2) is now a matter of taking appropriate direct products, since

$$\chi^{(m)}(\varphi)\chi^{(m')}(\varphi) = \chi^{(m+m')}(\varphi)\,.$$

In particular, the m-fold products $(x \pm iy)^m$ generate irreducible representations for the m-fold direct product, as discussed in Sec. 6.5. This

can be verified directly from the transformation (8.9) applied to x and y:

$$\begin{aligned} x' &= x\cos\varphi - y\sin\varphi\,, \\ y' &= x\sin\varphi + y\cos\varphi\,. \end{aligned}$$

Then,

$$\begin{aligned} (x' \pm iy')^m &= \Big[x\cos\varphi - y\sin\varphi \pm i(x\sin\varphi + y\cos\varphi)\Big]^m \\ &= \Big[x(\cos\varphi \pm i\sin\varphi) \pm iy(\cos\varphi \pm i\sin\varphi)\Big]^m \\ &= \Big[(x \pm iy)\,\mathrm{e}^{\pm i\varphi}\Big]^m \\ &= (x \pm iy)^m\,\mathrm{e}^{\pm im\varphi}\,. \end{aligned}$$

Therefore, we can now complete the character table for SO(2), including the basis functions which generate the irreducible representations:

SO(2)	E	$R(\varphi)$
$\Gamma^{\pm m}$: $(x \pm iy)^m$	1	$\mathrm{e}^{\pm im\varphi}$

We note for future reference that the basis functions $(x \pm iy)^m$ could have been derived in a completely different manner. Consider Laplace's equation in two dimensions:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0\,.$$

This equation is invariant under all the elements of SO(2), as can be easily verified. The general solution to this equation is

$$u(x, y) = F(x + iy) + G(x - iy)\,,$$

where F and G are arbitrary functions. Thus, if we are interested in solutions which are homogeneous polynomials of degree m, we can

choose in turn solutions with $F(s) = s^m$ and $G(s) = 0$ and then with $F(s) = 0$ and $G(s) = s^m$. We thereby obtain the expressions

$$u(x, y) = (x \pm iy)^m \tag{8.11}$$

as solutions of Laplace's equations which are also the basis functions of the irreducible representations of SO(2). These functions are the analogues in two dimensions of spherical harmonics, which are the solutions of Laplace's equations in three dimensions. These will be discussed later in this chapter.

8.3 Axis–Angle Representation of Proper Rotations in Three Dimensions

The three most common parametrizations of proper rotations were discussed in Section 7.4. For the purposes of obtaining the orthogonality relations for the characters of SO(3), the representation in terms of a fixed axis about which a rotation is carried out—the axis–angle representation—is the most convenient. We begin this section by showing how this representation emerges naturally from the basic properties of orthogonal matrices.

8.3.1 Eigenvalues of Orthogonal Matrices

Let A be any proper rotation matrix in three dimensions. Denoting the eigenvalues of A by λ_1, λ_2, and λ_3, and the corresponding eigenvalues by $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{u}_3$, we then have

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$$

for $i = 1, 2, 3$. We can also form the adjoint of each equation:

$$\boldsymbol{u}_i^\dagger A^{\mathrm{t}} = \lambda^* \boldsymbol{u}_i^\dagger \,.$$

These eigenvalue equations imply

$$\boldsymbol{u}_i^\dagger \boldsymbol{u}_i = \boldsymbol{u}_i^\dagger A^{\mathrm{t}} A \boldsymbol{u}_i = |\lambda_i^2| \boldsymbol{u}_i^\dagger \boldsymbol{u}_i \,,$$

which shows that $|\lambda_i^2| = 1$, i.e., that the modulus of every eigenvalue of an orthogonal matrix is unity [cf. (8.10)]. The most general form of such a quantity is a complex number of the form $e^{i\varphi}$ for some angle φ. But these eigenvalues are also the roots of the characteristic equation $\det(A - \lambda I) = 0$ so, according to the Fundamental Theorem of Algebra,[3] if they are complex, they must occur in complex conjugate pairs (because the coefficients of this polynomial, which are obtained from the entries of A, are real). Hence, the most general form of the eigenvalues of an orthogonal matrix in three dimensions is

$$\lambda_1 = 1, \qquad \lambda_2 = e^{i\varphi}, \qquad \lambda_3 = e^{-i\varphi}. \tag{8.12}$$

The eigenvector corresponding to $\lambda_1 = 1$, which is unaffected by the action of A, thereby defines the axis about which the rotation is taken. The quantity φ appearing in λ_2 and λ_3 defines the angle of rotation about this axis.

8.3.2 The Axis and Angle of an Orthogonal Matrix

In this section, we show how the axis and angle of an orthogonal matrix can be determined from its matrix elements. We take the axis of the rotation to be a unit vector $\boldsymbol{n}$, which is the eigenvector corresponding to the eigenvalue of unity:

$$A\boldsymbol{n} = \boldsymbol{n}. \tag{8.13}$$

This equation and the orthogonality of A ($AA^t = A^tA = 1$) enables us to write

$$A^t\boldsymbol{n} = A^tA\boldsymbol{n} = \boldsymbol{n}. \tag{8.14}$$

Subtracting (8.14) from (8.13) yields

$$(A - A^t)\boldsymbol{n} = 0.$$

[3] K. Hoffman and R. Kunze, *Linear Algebra* 2nd edn (Prentice–Hall, Englewood Cliffs, NJ, 1971), p. 138.

In terms of the matrix elements a_{ij} of A and the components n_i of $\boldsymbol{n}$, we then have

$$
\begin{aligned}
(a_{12} - a_{21})n_2 + (a_{13} - a_{31})n_3 &= 0\,,\\
(a_{21} - a_{12})n_1 + (a_{23} - a_{32})n_3 &= 0\,,\\
(a_{31} - a_{13})n_1 + (a_{32} - a_{23})n_2 &= 0\,.
\end{aligned}
$$

Notice that these equations involve only the *off-diagonal* elements of A. The solution of these equations yield the relations

$$\frac{n_2}{n_1} = \frac{a_{31} - a_{13}}{a_{23} - a_{32}}, \qquad \frac{n_3}{n_1} = \frac{a_{12} - a_{21}}{a_{23} - a_{32}}\,, \tag{8.15}$$

which, when combined with the normalization condition

$$\boldsymbol{n}\cdot\boldsymbol{n} = n_1^2 + n_2^2 + n_3^2 = 1$$

determines $\boldsymbol{n}$ uniquely.

The angle of the rotation can be determined from the invariance of the trace of A under similarity transformations. Noting that the trace is the sum of the eigenvalues, and using (8.12), we have

$$a_{11} + a_{22} + a_{33} = 1 + \mathrm{e}^{i\varphi} + \mathrm{e}^{-i\varphi} = 1 + 2\cos\varphi\,, \tag{8.16}$$

so φ is determined only by the *diagonal* elements of A.

8.3.3 Normal Form of an Orthogonal Matrix

We conclude this section by deriving the form of a rotation matrix in an orthogonal coordinate system which naturally manifests the axis and angle. The diagonal form of a rotation matrix is clearly given by

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{e}^{i\varphi} & 0 \\ 0 & 0 & \mathrm{e}^{i\varphi} \end{pmatrix}.$$

The eigenvector $\boldsymbol{n}$ corresponding to $\lambda_1 = 1$ is the axis of the rotation and can always be chosen to be real. However, the eigenvectors of $\lambda_2 =$

$e^{i\varphi}$ and $\lambda_3 = e^{-i\varphi}$ are inherently complex quantities. An orthonormal set can be chosen as

$$\boldsymbol{n}_2 = \tfrac{1}{2}\sqrt{2}(0,1,i), \qquad \boldsymbol{n}_2 = \tfrac{1}{2}\sqrt{2}(0,1,-i)\,,$$

respectively. Since we are interested in transformations of real coordinates, we must perform a unitary transformation from this complex basis to a real orthogonal basis, in which case our rotation matrix Λ will no longer be diagonal. The required unitary matrix which accomplishes this is

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \tfrac{1}{2}\sqrt{2} & \tfrac{1}{2}i\sqrt{2} \\ 0 & \tfrac{1}{2}\sqrt{2} & -\tfrac{1}{2}i\sqrt{2} \end{pmatrix}.$$

Thus,

$$R = U^{-1}\Lambda U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}. \tag{8.17}$$

When expressed in this basis, the rotation matrix clearly displays the axis of rotation through the entry $R_{11} = 1$, and the angle of rotation through a 2×2 rotational submatrix in a plane perpendicular to this axis.

8.3.4 Parameter Space for SO(3)

The axis-angle representation of three-dimensional rotations provides a convenient parametrization of all elements of SO(3). We have seen that every element of SO(3) can be represented by a unit vector corresponding to the rotation axis and a scalar corresponding to the rotation angle. Thus, consider the space defined by the three quantities

$$(n_1\varphi, n_2\varphi, n_3\varphi)\,, \tag{8.18}$$

where $n_1^2 + n_2^2 + n_3^2 = 1$. Every direction is represented by a point on the unit sphere. Thus, defining an azimuthal angle ϕ and a polar angle

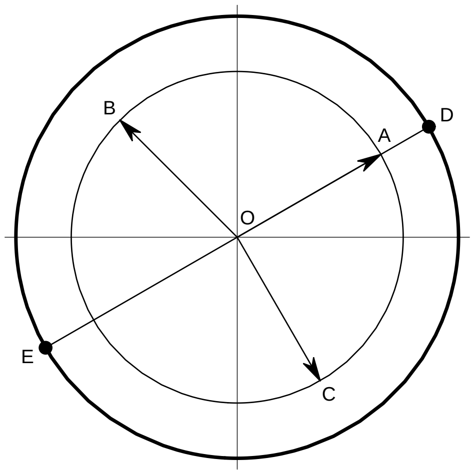


Figure 8.1: Two-dimensional representation of the parameter space of SO(3) as the interior of a sphere of radius π. The point A represents a rotation whose axis is along the direction OA and whose angle is the length of OA. The points at A, B and C correspond to rotations with the same angle but about axis along different directions. This defines the classes of SO(3). The diametrically opposite points at D and E correspond to the same operation.

θ according to the usual conventions in spherical polar coordinates, the parameter space of SO(3) can be represented as

$$(\varphi\cos\phi\sin\theta, \varphi\sin\phi\sin\theta, \varphi\cos\theta)\,, \tag{8.19}$$

where

$$0 \le \varphi \le \pi, \qquad 0 \le \phi \le 2\pi, \qquad 0 \le \theta \le \pi\,.$$

We can now see directly that this parameter space corresponds to the interior of a sphere of radius π (Fig. 8.1). For every point *within* the sphere, there is a unique assignment to an element of SO(3): the *direction* from the radius to the point corresponds to the direction of the rotation axis and the *distance* from the point to the origin represents the rotation angle. Two diametrically opposed points on the surface of the sphere ($\varphi = \pi$) correspond to the same rotation, since a rotation by π about $\boldsymbol{n}$ is the same as a rotation by $-\pi$ about this axis which, in turn, is the same as a rotation by π about $-\boldsymbol{n}$ (whatever the sense of rotation).

Another useful feature of the axis-angle parametrization is the representation of classes of SO(3). Consider two elements of SO(3) which have the same angle of rotation φ but about different axes $\boldsymbol{n}$ and $\boldsymbol{n}'$. We denote these operations by $R(\boldsymbol{n},\varphi)$ and $R(\boldsymbol{n}',\varphi)$. Let $U(\boldsymbol{n},\boldsymbol{n}')$ denote the rotation of $\boldsymbol{n}$ into $\boldsymbol{n}'$. The inverse of this operation then rotates $\boldsymbol{n}'$ into $\boldsymbol{n}$. The relationship between $R(\boldsymbol{n},\varphi)$, $R(\boldsymbol{n}',\varphi)$, and $U(\boldsymbol{n},\boldsymbol{n}')$ is, therefore,

$$R(\boldsymbol{n},\varphi) = \Big[U(\boldsymbol{n},\boldsymbol{n}')\Big]^{-1} R(\boldsymbol{n}',\varphi)U(\boldsymbol{n},\boldsymbol{n}')\,,$$

i.e., $R(\boldsymbol{n},\varphi)$ and $R(\boldsymbol{n}',\varphi)$ are related by a similarity transformation and, therefore, belong to the *same equivalence class*. Referring to Fig. 8.1, equivalence classes of SO(3) correspond to operations which lie on the same radius. Thus, *a summation over the classes of SO(3) is equivalent to a sum over spherical shells.*

8.4 Orthogonality Relations for SO(3)

The axis-angle representation of rotations provides, in addition to a conceptual simplicity of elements of SO(3) in parameter space, a natural framework within which to discuss the integration over the elements of SO(3) and thereby to obtain the Rearrangement Theorem for this group. In this section, we derive the density function g in (8.3) for this group and then use this to identify the appropriate form of the orthogonality relations for characters

8.4.1 The Density Function

As discussed in the introduction, one of the basic quantities of interest for continuous groups is the density of group elements as a function of position in parameter space. To determine this function for SO(3), we first consider the elements in the neighborhood of the identity and then examine the behavior of these points under an arbitrary element of SO(3). Referring to the discussion in Section 7.4.2, these elements correspond to rotations by infinitesimal angles φ_1, φ_2, and φ_3 about each of the three coordinate axes. The rotation matrix associated with

this transformation is

$$\delta R = \begin{pmatrix} 1 & -\varphi_3 & \varphi_2 \\ \varphi_3 & 1 & -\varphi_1 \\ -\varphi_2 & \varphi_1 & 1 \end{pmatrix} .$$

The identity of SO(3) corresponds to the origin in the three-dimensional parameter space, $\varphi_1 = \varphi_2 = \varphi_3 = 0$, and is indicated by the point O in Fig. 8.1. For infinitesimal rotation angles, the parameter space spanned by δR is associated with an infinitesimal volume element in the neighborhood of the origin.

We now follow the infinitesimal transformation δR by a *finite* transformation $R(\boldsymbol{n}, \varphi)$, i.e., we form the product $R\,\delta R$. This generates a volume element in the neighborhood of R and the product $R\,\delta R$ can be viewed as transformation of the volume near the origin to that near R. The Jacobian of this transformation is the relative change of volume near the origin to that near R or, equivalently, is the relative change of the *density* of operations near the origin to that near R. According to the discussion in the introduction, this is the information required from the density function for SO(3).

We have already seen that equivalence classes of SO(3) are comprised of all rotations with the same rotation angle, regardless of the direction of the rotation axis. Thus, the density function is expected to depend only on φ. Referring to Fig. 8.1, this means that the density of elements depends only on the "radial" distance from the origin, not on the direction, so we can choose R in accordance with this at our convenience. Therefore, in constructing the matrix $R\,\delta R$, we will use for R a matrix of the form in (8.17). Thus,

$$R\,\delta R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} 1 & -\varphi_3 & \varphi_2 \\ \varphi_3 & 1 & -\varphi_1 \\ -\varphi_2 & \varphi_1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -\varphi_3 & \varphi_2 \\ \varphi_3\cos\varphi + \varphi_2\sin\varphi & \cos\varphi - \varphi_1\sin\varphi & -\varphi_1\cos\varphi - \sin\varphi \\ \varphi_3\sin\varphi - \varphi_2\cos\varphi & \sin\varphi - \varphi_1\cos\varphi & -\varphi_1\sin\varphi + \cos\varphi \end{pmatrix} .$$

We can now use (8.15) and (8.16) to determine the axis $\boldsymbol{n}'$ and angle φ' of this product. The angle is determined from

$$1 + 2\cos\varphi' = 1 + 2\cos\varphi - 2\varphi_1 \sin\varphi\,,$$

which, upon cancelling common factors, becomes

$$\cos\varphi' = \cos\varphi - \varphi_1 \sin\varphi\,.$$

Using the standard trigonometric formula for the cosine of a sum, we find, to first order in φ_1, that

$$\varphi' = \varphi + \varphi_1\,.$$

The *unnormalized* components of $\boldsymbol{n}'$ are determined from (8.15) to be

$$\begin{aligned}
n_1' &= -2\varphi_1 \cos\varphi - 2\sin\varphi\,,\\
n_2' &= \varphi_3 \sin\varphi - \varphi_2(1+\cos\varphi)\,,\\
n_3' &= -\varphi_2 \sin\varphi - \varphi_3(1+\cos\varphi)\,.
\end{aligned}$$

To normalize the axis, we first determine the length based on these components. To first order in the φ_i, we find

$$|\boldsymbol{n}'| = 2\varphi_1 \cos\varphi + 2\sin\varphi\,.$$

Thus, the components of the normalized rotation axis of $R\,\delta R$ are

$$\begin{aligned}
n_1' &= 1\,,\\
n_2' &= -\tfrac{1}{2}\varphi_3 + \tfrac{1}{2}\varphi_2 \frac{1+\cos\varphi}{\sin\varphi}\,,\\
n_3' &= \tfrac{1}{2}\varphi_2 + \tfrac{1}{2}\varphi_3 \frac{1+\cos\varphi}{\sin\varphi}\,.
\end{aligned}$$

Expressed in terms of the parametrization in (8.18), $R\,\delta R$ is given by

$$(n_1'\varphi', n_2'\varphi', n_3'\varphi') =$$
$$\left\{\varphi + \varphi_1, \tfrac{1}{2}\varphi\left(-\varphi_3 + \varphi_2\frac{1+\cos\varphi}{\sin\varphi}\right), \tfrac{1}{2}\varphi\left(\varphi_2 + \varphi_3\frac{1+\cos\varphi}{\sin\varphi}\right)\right\}.$$

This defines the transformation from the neighborhood of the origin to the neighborhood near $R\,\delta R$. The Jacobian J of this transformation, obtained from

$$J = \det\left|\frac{\partial(n_i'\varphi')}{\partial\varphi_j}\right|, \tag{8.20}$$

determines how the density of elements of SO(3) near the origin is transformed to the density of points near R. By taking the derivatives in (8.20) to obtain the entries (i,j) in the Jacobian matrix, we obtain

$$J = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \varphi\dfrac{1+\cos\varphi}{2\sin\varphi} & -\frac{1}{2}\varphi \\ 0 & \frac{1}{2}\varphi & \varphi\dfrac{1+\cos\varphi}{2\sin\varphi} \end{vmatrix} = \frac{\varphi^2}{2(1-\cos\varphi)}\,.$$

Notice that

$$\lim_{\varphi\to 0} J = 1\,,$$

so that the normalization of the volume in parameter space is such that the volume near the unit element is unity. Hence, the *density* of elements in parameter space is the *reciprocal* of J, so the density function g in (8.3) is

$$g(\varphi) = \frac{2}{\varphi^2}(1-\cos\varphi)\,. \tag{8.21}$$

8.4.2 Integrals in Parameter Space

The density function in (8.21) now permits us to carry out integral over the group. Thus, for a function $F(\varphi,\Omega)$, where Ω denotes the angular variables in the parametrization in (8.19), we have

$$\iint g(\varphi)F(\varphi,\Omega)\varphi^2\,\mathrm{d}\varphi\,\mathrm{d}\Omega\,,$$

where we have used the usual volume element for spherical polar coordinates. Using the density function in (8.21), this integral becomes

$$\iint 2(1-\cos\varphi)F(\varphi,\Omega)\mathrm{d}\varphi\,\mathrm{d}\Omega\,.$$

We can now establish the orthogonality relation for characters. If we denote the characters for two irreducible representations of SO(3) by $\chi^\mu(\varphi)$ and $\chi^\nu(\varphi)$, then we have

$$\iint 2(1-\cos\varphi)\chi^\mu(\varphi)\chi^\nu(\varphi)\mathrm{d}\varphi\,\mathrm{d}\Omega = \delta_{\mu,\nu}\iint 2(1-\cos\varphi)\mathrm{d}\varphi\,\mathrm{d}\Omega\,.$$

The integral on the right-hand side of this equation, which has the value $8\pi^2$, corresponds to the volume of SO(3) in parameter space. The integral over the angular variables on the left-hand side yields $2\times 4\pi$, so cancelling common factors, we obtain

$$\int_0^\pi (1-\cos\varphi)\chi^\mu(\varphi)\chi^\nu(\varphi)\,\mathrm{d}\varphi = \pi\delta_{\mu,\nu}\,. \tag{8.22}$$

This is the orthogonality relation for characters of SO(3).

8.5 Irreducible Representations and Characters for SO(3)

For SO(2), we were able to determine the characters of the irreducible representations directly, i.e., without having to determine the basis functions of these representations. The structure of SO(3), however, does not allow for such a simple procedure, so we must determine the basis functions from the outset.

8.5.1 Spherical Harmonics

We proceed as in Section 8.2 by determining the homogeneous polynomial solutions of Laplace's equation, now in three dimensions:

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=0\,.$$

We seek solutions of the form

$$u(x,y,z)=\sum_{a,b} c_{ab}(x+iy)^a(x-iy)^b z^{\ell-a-b}\,,$$

which are homogeneous polynomials of degree ℓ. In spherical polar coordinates,

$$x = r\cos\phi\sin\theta, \qquad y = r\sin\phi\sin\theta, \qquad z = r\cos\theta\,,$$

where $0 \le \phi < 2\pi$ and $0 \le \theta \le \pi$, these polynomial solutions transform to

$$u(r,\theta,\phi) = \sum_{a,b} c_{ab} r^\ell \sin^{a+b}\theta\cos^{\ell-a-b}\theta\,\mathrm{e}^{i(a-b)\phi}\,. \tag{8.23}$$

Alternatively, Laplace's equation in spherical polar coordinates is

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\phi^2} = 0\,.$$

When the method of separation of variables is used to find solutions of this equation of the form $u(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi)$, the stipulation that the solution be single-valued with respect changes in ϕ by 2π,

$$u(r,\theta,\phi+2\pi) = u(r,\theta,\phi)\,,$$

requires that

$$\Phi(\phi) \propto \mathrm{e}^{im\phi}\,,$$

where m is an integer. Comparing this expression with the corresponding factor in (8.23), we see that $a - b = m$. Since the ranges of both a and b are between 0 and ℓ, we see that there are $2\ell + 1$ values of m consistent with homogeneous polynomial solutions of degree ℓ. The corresponding values of m are $-\ell \le m \le \ell$. The $2\ell + 1$ independent homogeneous polynomials of degree ℓ are called the **spherical harmonics** and denoted by $Y_{\ell m}(\theta,\phi)$. Their functional form is

$$Y_{\ell m}(\theta,\phi) \propto P_m^\ell(\theta)\,\mathrm{e}^{im\phi}\,, \tag{8.24}$$

where $P_m^\ell(\theta)$ is a **Legendre function**. In the following discussion, we will utilize only the exponential factor in the spherical harmonics.

8.5.2 Characters of Irreducible Representations

The $Y_{\ell m}(\theta,\phi)$ form a $(2\ell+1)$-dimensional representation of SO(3). Thus, for a general rotation R, we have

$$R\,Y_{\ell m}(\theta,\phi) = \sum_{m'=-\ell}^{\ell} Y_{\ell m'}(\theta,\phi)\,.\Gamma^{\ell}_{m'm}(R)$$

To determine the character of this representation, it is convenient to again invoke the fact that the classes of SO(3) are determined only by the rotation angle, not by the direction of the rotation axis. Thus, we can choose a rotation axis at our convenience and we therefore focus on rotations through an angle φ about the z-axis. In this case, the form of (8.24) allows us to write

$$R_z(\varphi)Y_{\ell m}(\theta,\phi) = Y_{\ell m}(\theta,\phi-\varphi) = \mathrm{e}^{-im\varphi}Y_{\ell m}(\theta,\phi)\,.$$

Thus, the corresponding transformation matrix is given by

$$\Gamma^{\ell}[R_z(\varphi)] = \begin{pmatrix} \mathrm{e}^{-i\ell\varphi} & 0 & \cdots & 0 \\ 0 & \mathrm{e}^{-i(\ell-1)\varphi} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathrm{e}^{i\ell\varphi} \end{pmatrix}. \tag{8.25}$$

The character $\chi^{(\ell)}(\varphi)$ of this class is obtained by taking the trace of this matrix:

$$\begin{aligned}
\chi^{(\ell)}(\varphi) &= \mathrm{e}^{-i\ell\varphi} + \mathrm{e}^{-i(\ell-1)\varphi} + \cdots + \mathrm{e}^{i\ell\varphi} \\
&= \mathrm{e}^{-i\ell\varphi}\Big(1 + \mathrm{e}^{i\varphi} + \mathrm{e}^{2i\varphi} + \cdots + \mathrm{e}^{2\ell i\varphi}\Big) \\
&= \mathrm{e}^{-i\ell\varphi}\frac{1-\mathrm{e}^{-(2\ell+1)i\varphi}}{1-\mathrm{e}^{i\varphi}} \\
&= \frac{\mathrm{e}^{(\ell+1/2)i\varphi} - \mathrm{e}^{-(\ell+1/2)i\varphi}}{\mathrm{e}^{i\varphi/2} - \mathrm{e}^{-i\varphi/2}} \\
&= \frac{\sin\left[(\ell+\frac{1}{2})\varphi\right]}{\sin\left(\frac{1}{2}\varphi\right)}\,.
\end{aligned}$$

The orthogonality integral for these characters takes the form

$$\int_0^\pi (1-\cos\varphi)\frac{\sin\left[(\ell+\frac{1}{2})\varphi\right]\sin\left[(\ell'+\frac{1}{2})\varphi\right]}{\sin^2\left(\frac{1}{2}\varphi\right)}\,\mathrm{d}\varphi\,.$$

Using the trigonometric identity

$$2\sin^2\left(\tfrac{1}{2}\varphi\right) = 1-\cos\varphi$$

enables us to write the orthogonality integral as

$$\int_0^\pi \sin\left[(\ell+\tfrac{1}{2})\varphi\right]\sin\left[(\ell'+\tfrac{1}{2})\varphi\right]\mathrm{d}\varphi = \tfrac{1}{2}\pi\delta_{\ell,\ell'}\,,$$

where the right-hand side of this equation follows either from (8.22) or from the orthogonality of the sine functions over $(0,\pi)$.

It is possible to show directly, using Schur's first lemma, that the spherical harmonics form a basis for $(2\ell+1)$-dimensional *irreducible* representations of SO(3). However, this requires invoking properties of the Legendre functions in (8.24). If we confine ourselves to the matrices in (8.25) then we can show that a matrix that commutes with all such rotation matrices must reduce to a diagonal matrix. If we then consider rotations about any other direction, which requires some knowledge of the Legendre functions, we can then show that this constant matrix must, in fact, be a constant multiple of the unit matrix. Hence, according to Schur's first lemma, these representations are irreducible. We can now construct the character table for SO(3) with the basis functions which generate the irreducible representations:

SO(3)	E	$R(\varphi)$
Γ^ℓ: $Y_{\ell m}(\theta,\phi)$	1	$\dfrac{\sin\left[(\ell+\frac{1}{2})\varphi\right]}{\sin\left(\frac{1}{2}\varphi\right)}$

8.6 Summary

In this chapter, we have shown how the orthogonality relations developed for finite groups must be adapted for continuous groups, using

SO(2) and SO(3) as examples. For SO(2), which is a one-parameter Abelian group, this proved to be a straightforward matter. However, the corresponding calculations for SO(3) required us to determine explicitly the density function to produce the appropriate form of the orthogonality relations. We found that the there are an infinite sequence of irreducible representations of dimensionality $2\ell + 1$, where $\ell \geq 0$. Because of the connection between SO(3) and angular momentum, the structure of these irreducible representations has several physical consequences:

- For systems that possess spherical symmetry, the energy eigenstates have degeneracies of $2\ell + 1$. The fact that there is a greater degeneracy for the hydrogen atom is due to a "hidden" SO(4) symmetry.[4]

- The formation and decomposition of direct products of the irreducible representations of SO(3) forms the basis of angular momentum coupling rules (Problem 6, Problem Sets 10) and the classification of atomic spectra.[5]

- When atoms are placed within crystals, the original spherical symmetry is lowered to the symmetry of the crystal. This causes levels which were degenerate in the spherically-symmetric environment to split. Such "crystal-field" effects are important for many aspects for electrons in crystalline solids.[6]

---

[4] H.F. Jones, *Groups, Representations and Physics* (Institute of Physics, Bristol, 1998), pp. 124–127.

[5] E.P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959), pp. 177–194.

[6] M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw–Hill, New York, 1964), pp. 65–80.

Chapter 9

Unitary Groups and SU(N)*

The irreducible representations of SO(3) are appropriate for describing the degeneracies of states of quantum mechanical systems which have rotational symmetry in three dimensions. But there are many systems for which operations on classical coordinates must be supplemented by operations on "internal" degrees of freedom which have no classical analogue. For example, the Stern–Gerlach experiment showed that electrons are endowed with an internal degree of freedom called "spin" which has the properties of an angular momentum. The two spin states are therefore inconsistent with the dimensionalities of the irreducible representations of SO(3), so another group—SU(2)—must be used to describe these states. Since, as we will show in Section 9.2, SU(2) is locally isomorphic to SO(3), we can define a total spin S in an abstract three-dimensional space, analogous to the total angular momentum in real space. In particle physics, unitary symmetry was used to describe the approximate symmetry (called isospin) of neutrons and protons and, more recently, to describe particle spectra within the framework of the quark model.

In this chapter, we introduce unitary groups and their irreducible representations in a similar manner to which we developed SO(3). We begin by defining unitarity in terms of the invariance of an appropriate quantity and proceed to discuss the construction of irreducible representations of these groups in N dimensions. Higher-dimensional irreducible representations will be obtained with the aid of Young tableaux, which is a diagrammatic technique for determining the dimensionali-

ties and the basis functions of irreducible representations derived from direct products.

9.1 SU(2)

As with orthogonal matrices, the unitary groups can be defined in terms of quantities which are left invariant. Consider a general complex transformation in two dimensions, $\boldsymbol{x}' = A\boldsymbol{x}$ which, in matrix form, reads:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

where a, b, c, and d are complex, so there are eight free parameters. The determinant of this matrix is nonzero to permit the construction of inverses.

9.1.1 Unitary Transformations

Suppose we require the quantity $|x|^2 + |y|^2$ to be an invariant of such a transformation. Then,

$$\begin{aligned} |x'|^2 + |y'|^2 &= |ax + by|^2 + |cx + dy|^2 \\ &= (ax + by)(a^*x^* + b^*y^*) + (cx + dy)(c^*x^* + d^*y^*) \\ &= (|a|^2 + |c|^2)|x|^2 + (ab^* + cd^*)xy^* + (a^*b + c^*d)x^*y \\ &\quad + (|b|^2 + |d|^2)|y|^2 \\ &= |x|^2 + |y|^2 \end{aligned}$$

Since x and y are independent variables, this invariance necessitates setting the following conditions on the matrix elements:

$$|a|^2 + |c|^2 = 1, \qquad |b|^2 + |d|^2 = 1, \qquad ab^* + cd^* = 0$$

These four conditions (the last equation provides two conditions because it involves complex quantities) means that the original eight free

parameters are reduced to four. These conditions are the same as those obtained by requiring the $A^\dagger A = 1$, so the determinant of the resulting matrix has modulus unity. These transformations are analogous to orthogonal transformations of real coordinates and, indeed, orthogonal transformations are also unitary. The group comprised of unitary matrices is denoted by U(2) and by U(N) for the N-dimensional case.

9.1.2 Special Unitary Transformations

If, in addition to the conditions above, we require that the determinant of the transformation is unity, the transformation matrix must have the form

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \qquad |a|^2 + |b|^2 = 1 \tag{9.1}$$

There are now three free parameters and the group of these matrices is denoted by SU(2) where, as in our discussion of orthogonal groups, the 'S' signifies 'special' because of the requirement of a unit determinant.

9.2 Relation between SU(2) and SO(3)

9.2.1 Pauli Matrices

If the matrix elements of the general unitary matrix in (9.1) are expressed in terms of their real and imaginary parts, we can decompose this matrix into the components of a "basis." Thus, with $a = a_r + ia_i$ and $b = b_r + ib_i$, we have

$$\begin{aligned} U &= \begin{pmatrix} a_r + ia_i & b_r + ib_i \\ -b_r + ib_i & a_r - ia_i \end{pmatrix} \\ &= a_r \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + ia_i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \underbrace{b_r \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{ib_r \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}} + ib_i \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Thus, any 2×2 unitary matrix can be represented as a linear combination of the unit matrix and the matrices

$$\sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

These three (Hermitian) matrices are known as the **Pauli matrices**. They satisfy the following multiplication rules:

$$\begin{aligned} &\sigma_i^2 = I && (i = x, y, z) \\ &\sigma_i \sigma_j = -\sigma_j \sigma_i = i\varepsilon_{ijk}\sigma_k && (\{i, j, k\} = x, y, z) \end{aligned} \tag{9.2}$$

where I is the 2×2 unit matrix. These multiplication rules can be used to obtain a concise expression for the product of two matrices written as $\boldsymbol{a}\cdot\boldsymbol{\sigma}$ and $\boldsymbol{b}\cdot\boldsymbol{\sigma}$, where $\boldsymbol{a} = (a_x, a_y, a_z)$, $\boldsymbol{b} = (b_x, b_y, b_z)$, and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$:

$$(\boldsymbol{a}\cdot\boldsymbol{\sigma})(\boldsymbol{b}\cdot\boldsymbol{\sigma}) = (\boldsymbol{a}\cdot\boldsymbol{b})I + i(\boldsymbol{a} \times \boldsymbol{b})\cdot\boldsymbol{\sigma} \tag{9.3}$$

9.2.2 Infinitesimal Generators

Moreover, if we define matrices $X_i = -\frac{1}{2}i\sigma_i$, for $i = 1, 2, 3$, then the second of the multiplication rules in (9.2) yield the following commutation relations:

$$[X_i, X_j] = \varepsilon_{ijk}X_k$$

These are identical to commutators of the infinitesimal generators of SO(3) in (7.13). Thus, locally at least, there is an isomorphism between SO(3) and SU(2). Motivated by the discussion in Section 7.3, consider the matrix

$$U = \exp\left(- \tfrac{1}{2}i\varphi\boldsymbol{n}\cdot\boldsymbol{\sigma}\right)$$

where $\varphi\boldsymbol{n}$ is the axis-angle representation of a rotation (Section (8.3). Since the exponential of a matrix is defined by its Taylor series expansion, we have

$$U = \sum_{k=0}^{\infty} \frac{(-i)^n}{n!}(\tfrac{1}{2}\varphi)^n(\boldsymbol{n}\cdot\boldsymbol{\sigma})^n$$

From Equation (9.3), $(\boldsymbol{n}\cdot\boldsymbol{\sigma})^2 = I$, so

$$
\begin{aligned}
U &= I\sum_{k=0}^{\infty}\frac{(-1)^n}{(2n)!}\left(\tfrac{1}{2}\varphi\right)^{2n} - i(\boldsymbol{n}\cdot\boldsymbol{\sigma})\sum_{k=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\left(\tfrac{1}{2}\varphi\right)^{2n+1} \\
&= \cos\left(\tfrac{1}{2}\varphi\right)I - i(\boldsymbol{n}\cdot\boldsymbol{\sigma})\sin\left(\tfrac{1}{2}\varphi\right) \\
&= \begin{bmatrix} \cos\left(\tfrac{1}{2}\varphi\right) - in_z\sin\left(\tfrac{1}{2}\varphi\right) & -(n_y + in_x)\sin\left(\tfrac{1}{2}\varphi\right) \\ (n_y - in_x)\sin\left(\tfrac{1}{2}\varphi\right) & \cos\left(\tfrac{1}{2}\varphi\right) + in_z\sin\left(\tfrac{1}{2}\varphi\right) \end{bmatrix}
\end{aligned}
\tag{9.4}
$$

This matrix is manifestly of the unitary form in (9.1) with unit determinant. The Pauli matrices are, therefore, the infinitesimal generators of SU(2) and form a representation of its Lie algebra.

9.2.3 Local and Global Mappings between SU(2) and SO(3)

The matrix in (9.4) is parametrized in the same way as rotations in SO(3), namely, in terms of a rotation angle φ and a rotation axis $\boldsymbol{n}$. But, although the mapping between SU(2) and SO(3) is *locally* an isomorphism, since their algebras are isomorphic, *globally* this relationship is a homomorphism. The reason for this stems from the periodicity of the two groups: SO(3) has a periodicity of 2π, while SU(2) has a periodicity of 4π. In particular $U(0,\boldsymbol{n}) = I$, but $U(2\pi,\boldsymbol{n}) = -I$, so both of these elements are associated with the identity of SO(3). Moreover, these elements form an invariant subgroup of SU(2) (Section 2.4) which is isomorphic to the group $\mathrm{Z}_2 = \{1,-1\}$ under ordinary multiplication. In general, using the trigonometric identities,

$$
\begin{aligned}
\cos\left[\tfrac{1}{2}(\varphi+2\pi)\right] &= -\cos\left(\tfrac{1}{2}\varphi\right) \\
\sin\left[\tfrac{1}{2}(\varphi+2\pi)\right] &= -\sin\left(\tfrac{1}{2}\varphi\right)
\end{aligned}
$$

we find that

$$
U(\varphi+2\pi,\boldsymbol{n}) = -U(\varphi,\boldsymbol{n})
$$

Thus, if we form the cosets of the subgroup $\{U(0,\boldsymbol{n}),U(2\pi,\boldsymbol{n})\}$, we obtain

$$\Big\{U(0,\boldsymbol{n}),U(2\pi,\boldsymbol{n})\Big\}U(\varphi,\boldsymbol{n})=\Big\{U(\varphi,\boldsymbol{n}),U(\varphi+2\pi,\boldsymbol{n})\Big\}$$

Thus, the factor group $\mathrm{SU}(2)/\mathrm{Z}_2$ is isomorphic to SO(3):

$$\mathrm{SU}(2)/\mathrm{Z}_2=\mathrm{SO}(3)$$

In fact, this double-valuedness extends to characters as well. Taking the trace of the matrix in 9.4) yields

$$2\cos\left(\tfrac{1}{2}\varphi\right)$$

If we compare this expression with that for $\chi^{(\ell)}(\varphi)$ for SO(3) with $\ell=\frac{1}{2}$, we find

$$\chi^{(1/2)}(\varphi)=\frac{\sin\varphi}{\sin\left(\frac{1}{2}\varphi\right)}=2\cos\left(\tfrac{1}{2}\varphi\right)$$

so the two-dimensional (irreducible) representation of SU(2) generated by the Pauli matrices corresponds to a representation of SO(3) with a half-integer index. The integer values of ℓ can be traced to the requirement of *single-valuedness* of the spherical harmonics, so the double-valued correspondence between SU(2) and SO(3) results in this half-integer index.

9.3 Irreducible Representations of SU(2)

When we constructed the irreducible representations of SO(2) and SO(3), we used as basis functions obtained from the coordinates $\{x,y\}$ and $\{x,y,z\}$, respectively, and to obtain higher-order irreducible representations from direct products. The basic procedure is much the same for unitary groups, except that we can no longer rely on basis states expressed in terms of coordinates. In this section, we carry out the required calculations for SU(2) and then generalize the method for SU(N) in the next section.

9.3.1 Basis States

By associating the Pauli matrices with angular momentum operators through $J_i = \frac{1}{2}\hbar\sigma_i$, we choose as our basis states the vectors

$$u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

There are several physical interpretations of these states. For example, they can represent the two possible energy eigenstates of a spin-$\frac{1}{2}$ particle, such an electron or proton. Another possibility is that u_1 and u_2 represent the isospin eigenstates of an isospin-$\frac{1}{2}$ particle, such as a proton or a neutron. The fact that the proton and neutron are not exactly degenerate means that isospin symmetry is only an approximate symmetry. A third interpretation of u_1 and u_2 is as "up" and "down" quarks which make up nucleons. We will discuss further refinements of the quark model in the context of SU(N) later in this chapter.

9.3.2 Multiparticle Systems and Direct Products

When using basis states of SU(2) to construct multiparticle states through direct products, we must respect the indistinguishability of the particles. Thus, *measurable* properties of a quantum system cannot depend on the labelling of the particles, though wavefunctions, of course, need not obey this invariance. Consider a two-particle system, with particle '1' in state i and particle '2' in state j. The corresponding wavefunction is $\psi_{i,j}(1,2)$. We require that

$$|\psi_{i,j}(1,2)|^2 = |\psi_{i,j}(2,1)|^2$$

which implies that

$$\psi_{i,j}(2,1) = \mathrm{e}^{i\theta}\psi_{i,j}(1,2)$$

for some phase angle θ. Since a two-fold exchange restores the original labelling,

$$\psi_{i,j}(1,2) = \mathrm{e}^{i\theta}\psi_{i,j}(2,1) = \mathrm{e}^{2i\theta}\psi_{i,j}(1,2)$$

we must have that $e^{2i\theta} = 1$, or that $\theta = 0$ or $\theta = \pi$. In the first case, the wavefunction is *symmetric* under the interchange of particles,

$$\psi_{i,j}(2,1) = \psi_{i,j}(1,2)$$

while in the latter case, the wavefunction is *antisymmetric* under the interchange of particles,

$$\psi_{i,j}(2,1) = -\psi_{i,j}(1,2)$$

Consider now a two-particle system each of which occupy one of the states of SU(2). The basis of these two-particle states is comprised of $\{u_1u_1, u_1u_2, u_2u_1, u_2u_2\}$, where we have adopted the convention that the order of the states corresponds to the order of the particle coordinates, e.g., $u_1u_1 \equiv u_1(1)u_1(2)$. But not all of these states are symmetric or antisymmetric under the interchange of particles. Hence, we construct the new basis

$$\Big\{\underbrace{u_1u_1, u_1u_2 + u_2u_1, u_2u_2}_{\text{symmetric}}, \underbrace{u_1u_2 - u_2u_1}_{\text{antisymmetric}}\Big\} \tag{9.5}$$

We can compare this result with that obtained from the two-fold direct product representation of SU(2):

$$\begin{aligned}
\chi^{(1/2)}(\varphi)\chi^{(1/2)}(\varphi) &= \Big[2\cos\big(\tfrac{1}{2}\big)\Big]^2 \\
&= \Big(e^{i\varphi/2} + e^{-i\varphi/2}\Big)^2 \\
&= \Big(e^{i\varphi/2} + 1 + e^{-i\varphi/2}\Big) + 1 \\
&= \chi^{(1)}(\varphi) + \chi^{(0)}(\varphi)
\end{aligned}$$

we see that the three symmetric wavefunctions for a basis for the $\ell = 1$ irreducible representation of SO(3) and the antisymmetric wavefunctions transforms as the identical representation ($\ell = 0$) of SO(3). If we think of these as spin-$\frac{1}{2}$ particles, the symmetric state corresponds to a total spin $S = 1$, while the antisymmetric state corresponds to $S = 0$. We could proceed in this way to construct states for larger numbers of particles, but in the next section we introduce a technique which is far more efficient and which can be applied to other SU(N) groups, where the direct method described in this section becomes cumbersome.

9.3.3 Young Tableaux

Determining the dimensionalities of the irreducible representations of direct products of basis states of SU(N) is a problem which is encountered in several applications in physics and group theory. **Young tableaux** provide a diagrammatic method for carrying this out in a straightforward manner. In this section, we repeat the calculation in the preceding section to illustrate the method, and in the next section, we describe the general procedure for applying Young tableaux to SU(N).

The basic unit of a Young tableau is a 'box', shown below

$$\square$$

which denotes a basis state. If there is no entry in the box, then this tableau represents any state. An entry, signified by a number denotes one of the basis states in some reference order. Thus, for SU(2), we have

$$u_1 = \boxed{1} \qquad\qquad u_2 = \boxed{2}$$

The utility of Young tableaux centers around the construction of direct products. For the two-fold direct products of SU(2) in (9.5), there are two types of states, symmetric and antisymmetric. The Young tableau for a generic two-particle symmetric state is

$$\square\square$$

and the two-particle antisymmetric state is

$$\begin{array}{c}\square\\\square\end{array}$$

The Young tableaux for three symmetric states in (9.5) are

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 2 & 2 \\ \hline\end{array}$$

and that for the antisymmetric state is

$$\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline\end{array}$$

In the framework of Young tableaux, the two-fold direct product is written as

$$\begin{array}{|c|}\hline \\ \hline\end{array} \times \begin{array}{|c|}\hline \\ \hline\end{array} = \begin{array}{|c|c|}\hline & \\ \hline\end{array} + \begin{array}{|c|}\hline \\ \hline \\ \hline\end{array}$$

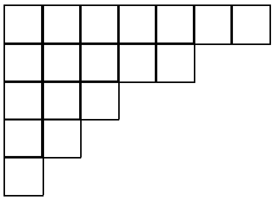
The three-fold direct product illustrates the conventions used in the construction of Young tableaux and their labelling. The generic tableaux are

$$\begin{array}{|c|}\hline \\ \hline\end{array} \times \begin{array}{|c|}\hline \\ \hline\end{array} \times \begin{array}{|c|}\hline \\ \hline\end{array} = \begin{array}{|c|c|c|}\hline & & \\ \hline\end{array} + \begin{array}{|c|c|}\hline & \\ \hline \\ \cline{1-1}\end{array} + \begin{array}{|c|}\hline \\ \hline \\ \hline \\ \hline\end{array}$$

The rules for constructing the "standard" arrangement of Young tableaux are as follows

- The rows are constructed from left to right
- The columns are constructed from top to bottom
- No row is longer than any row above it
- No column is longer than any column to the left of it

Thus, with these conventions, a typical tableau is shown below:



The states for the three-fold direct product are as follows. There are four symmetric states:

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 1 & 2 \\ \hline\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 2 & 2 \\ \hline\end{array} \qquad \begin{array}{|c|c|c|}\hline 2 & 2 & 2 \\ \hline\end{array}$$

which correspond to a four-dimensional irreducible representation, and two "mixed" states:

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

which correspond to a two-dimensional irreducible representation. There are no totally antisymmetric three-particle states because we have only two distinct basis states. Thus, the rules for entering states into Young tableaux are:

- The numbers within rows are *nondecreasing* from left to right.
- The numbers within columns are *increasing* from top to bottom.

The two sets of rules for constructing Young tableaux of generic states and identifying particular states enables the calculation of the dimensionalities in a straightforward manner, often by identifying appropriate combinatorial rules.

9.4 Young Tableaux for SU(N)

The groups SU(N) have acquired an importance in particle physics because of the quark model. This necessitates calculating direct products of basis states to determine the characteristics of particle spectra.

This, in turn, requires that we adapt the methodology of the Young tableaux developed in the preceding section to SU(N), which turns out to be straightforward given the rules stated in the preceding section. There is no change to the construction of the generic tableaux; the only changes are in the labelling of the tableaux. Consider, for example the case of a two-fold direct product of SU(3). There are six symmetric states

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 2 & 2 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 2 & 3 \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline 3 & 3 \\ \hline\end{array}$$

and three antisymmetric states

$$\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline\end{array} \qquad \begin{array}{|c|}\hline 1 \\ \hline 3 \\ \hline\end{array} \qquad \begin{array}{|c|}\hline 2 \\ \hline 3 \\ \hline\end{array}$$

As is evident from these constructions, the number of states associated with a tableau of a particular topology increases sharply with the number of basis states. The rules in the preceding section allow the number of such symmetric and antisymmetric states to be calculated for SU(N). There $\frac{1}{2}N(N+1)$ symmetric states and $\frac{1}{2}N(N-1)$ antisymmetric states.

The only other modification to our discussion of SU(2) is that for larger numbers of basis states, tableaux which make no contribution to SU(2), may make a contribution to SU(N). Consider, for example, the antisymmetric three-particle state. This state vanishes for SU(2) because there are only two basis states, but for SU(3), we have

$$\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline\end{array}$$

In fact, this is a direct consequence of the rule for labelling Young tableaux, and we see that, for SU(N), any column with more than N boxes makes no contribution.

9.5 Summary

In this chapter, we have extended our discussion of orthogonal groups to unitary groups. These groups play an especially important role in quantum mechanics because of their property of conserving probability density. We have constructed direct products of basis states, which are required in a number of applications of these groups. The use of Young tableaux was shown to be an especially convenient way to determine the dimensionalities of higher-dimensional irreducible representations of unitary groups and their basis functions.

Group Theory

Problem Set 1 October 12, 2001

Note: Problems marked with an asterisk are for Rapid Feedback; problems marked with a double asterisk are optional.

1. Show that the wave equation for the propagation of an impulse at the speed of light c,

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\,,$$

is covariant under the Lorentz transformation

$$x' = \gamma(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma\left(t - \frac{v}{c^2}x\right),$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$.

2.* The Schrödinger equation for a free particle of mass m is

$$i\hbar\frac{\partial \varphi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \varphi}{\partial x^2}\,.$$

Show that this equation is invariant to the global change of phase of the wavefunction:

$$\varphi \to \varphi' = \mathrm{e}^{i\alpha}\varphi\,,$$

where α is any real number. This is an example of an **internal** symmetry transformation, since it does not involve the space-time coordinates.

According to Noether's theorem, this symmetry implies the existence of a conservation law. Show that the quantity $\int_{-\infty}^{\infty} |\varphi(x,t)|^2\,dx$ is independent of time for solutions of the free-particle Schrödinger equation.

3.* Consider the following sets of elements and composition laws. Determine whether they are groups and, if not, identify which group property is violated.

(a) The rational numbers, excluding zero, under multiplication.

(b) The non-negative integers under addition.

(c) The even integers under addition.

(d) The nth roots of unity, i.e., $\mathrm{e}^{2\pi mi/n}$, for $m = 0, 1, \ldots, n-1$, under multiplication.

(e) The set of integers under ordinary *subtraction*.

4.** The general form of the Liouville equation is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}y}{\mathrm{d}x}\right] + \big[q(x) + \lambda r(x)\big]y = 0$$

where p, q and r are real-valued functions of x with p and r taking only positive values. The quantity λ is called the eigenvalue and the function y, called the eigenfunction, is assumed to be defined over an interval $[a, b]$. We take the boundary conditions to be

$$y(a) = y(b) = 0$$

but the result derived below is also valid for more general boundary conditions. Notice that the Liouville equations contains the one-dimensional Schrödinger equation as a special case.

Let $u(x;\lambda)$ and $v(x;\lambda)$ be the fundamental solutions of the Liouville equation, i.e. u and v are two linearly-independent solutions in terms of which all other solutions may be expressed (for a given value λ). Then there are constants A and B which allow any solution y to be expressed as a linear combination of this fundamental set:

$$y(x;\lambda) = Au(x;\lambda) + Bv(x;\lambda)$$

These constants are determined by requiring $y(x;\lambda)$ to satisfy the boundary conditions:

$$y(a;\lambda) = Au(a;\lambda) + Bv(a;\lambda) = 0$$

$$y(b;\lambda) = Au(b;\lambda) + Bv(b;\lambda) = 0$$

Use this to show that the solution $y(x;\lambda)$ is unique, i.e., that there is one and only one solution corresponding to an eigenvalue of the Liouville equation.

Group Theory

Solutions to Problem Set 1 October 6, 2000

1. To express the the wave equation for the propagation of an impulse at the speed of light c,

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\,, \tag{1}$$

under the Lorentz transformation

$$x'=\gamma(x-vt),\quad y'=y,\quad z'=z,\quad t'=\gamma\Big(t-\frac{v}{c^2}x\Big)\,, \tag{2}$$

where $\gamma=(1-v^2/c^2)^{-1/2}$, we need to obtain expressions for the second derivatives in the primed variables. With $u'(x',y',z')=u(x,y,t)$, we have

$$\frac{\partial u}{\partial x}=\frac{\partial u'}{\partial x'}\frac{\partial x'}{\partial x}+\frac{\partial u'}{\partial t'}\frac{\partial t'}{\partial x}=\gamma\frac{\partial u'}{\partial x'}-\gamma\frac{v}{c^2}\frac{\partial u'}{\partial t'}\,, \tag{3}$$

$$\frac{\partial^2 u}{\partial x^2}=\gamma^2\frac{\partial^2 u'}{\partial x'^2}+\gamma^2\frac{v^2}{c^4}\frac{\partial^2 u'}{\partial t'^2}\,, \tag{4}$$

$$\frac{\partial^2 u}{\partial y^2}=\frac{\partial^2 u'}{\partial y'^2}\,, \tag{5}$$

$$\frac{\partial^2 u}{\partial z^2}=\frac{\partial^2 u'}{\partial z'^2}\,, \tag{6}$$

$$\frac{\partial u}{\partial t}=\frac{\partial u'}{\partial t'}\frac{\partial t'}{\partial t}+\frac{\partial u'}{\partial x'}\frac{\partial x'}{\partial t}=\gamma\frac{\partial u'}{\partial t'}-\gamma v\frac{\partial u'}{\partial x'}\,, \tag{7}$$

$$\frac{\partial^2 u}{\partial t^2}=\gamma^2\frac{\partial^2 u'}{\partial t'^2}+\gamma^2 v^2\frac{\partial^2 u'}{\partial x'^2}\,. \tag{8}$$

Substituting these expressions into the wave equation yields

$$\frac{\gamma^2}{c^2}\frac{\partial^2 u'}{\partial t'^2}+\frac{\gamma^2 v^2}{c^2}\frac{\partial^2 u'}{\partial x'^2}=\gamma^2\frac{\partial^2 u'}{\partial x'^2}+\gamma^2\frac{v^2}{c^4}\frac{\partial^2 u'}{\partial t'^2}+\frac{\partial^2 u'}{\partial y'^2}+\frac{\partial^2 u'}{\partial z'^2}\,, \tag{9}$$

which, upon rearrangement, becomes

$$\frac{\gamma^2}{c^2}\left(1-\frac{v^2}{c^2}\right)\frac{\partial^2 u'}{\partial t'^2}=\gamma^2\left(1-\frac{v^2}{c^2}\right)\frac{\partial^2 u'}{\partial x'^2}+\frac{\partial^2 u'}{\partial y'^2}+\frac{\partial^2 u'}{\partial z'^2}\,. \tag{10}$$

Invoking the definition of γ, we obtain

$$\frac{1}{c^2}\frac{\partial^2 u'}{\partial t'^2}=\frac{\partial^2 u'}{\partial x'^2}+\frac{\partial^2 u'}{\partial y'^2}+\frac{\partial^2 u'}{\partial z'^2}\,, \tag{11}$$

which confirms the covariance of the wave equation under the Lorentz transformation.

2. We begin with the Schrödinger equation for a free particle of mass m:

$$i\hbar\frac{\partial\varphi}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2\varphi}{\partial x^2}\,. \tag{12}$$

Performing the transformation

$$\varphi\to\varphi'=\mathrm{e}^{i\alpha}\varphi\,, \tag{13}$$

where α is any real number, we find

$$i\hbar\mathrm{e}^{i\alpha}\frac{\partial\varphi'}{\partial t}=-\frac{\hbar^2}{2m}\mathrm{e}^{i\alpha}\frac{\partial^2\varphi'}{\partial x^2}\,, \tag{14}$$

or, upon cancelling the common factor $\mathrm{e}^{i\alpha}$,

$$i\hbar\frac{\partial\varphi'}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2\varphi'}{\partial x^2}\,, \tag{15}$$

which establishes the covariance of the Schrödinger equation under this transformation.

We can derive the corresponding quantity multiplying Eq. (12) by the complex conjugate φ^* and subtracting the product of φ and the complex conjugate of Eq. (12):

$$i\hbar\left(\varphi^*\frac{\partial\varphi}{\partial t}+\varphi\frac{\partial\varphi^*}{\partial t}\right)=-\frac{\hbar^2}{2m}\left(\varphi^*\frac{\partial^2\varphi}{\partial x^2}-\varphi\frac{\partial^2\varphi^*}{\partial x^2}\right). \tag{16}$$

The left-hand side of this equation can be written as

$$i\hbar\left(\varphi^*\frac{\partial\varphi}{\partial t}+\varphi\frac{\partial\varphi^*}{\partial t}\right)=i\hbar\frac{\partial}{\partial t}\left(\varphi\varphi^*\right)=i\hbar\frac{\partial}{\partial t}|\varphi|^2\,, \qquad (17)$$

and the right-hand side can be written as

$$-\frac{\hbar^2}{2m}\left(\varphi^*\frac{\partial^2\varphi}{\partial x^2}-\varphi\frac{\partial^2\varphi^*}{\partial x^2}\right)=-\frac{\hbar^2}{2m}\frac{\partial}{\partial x}\left(\varphi^*\frac{\partial\varphi}{\partial x}-\varphi\frac{\partial\varphi^*}{\partial x}\right). \qquad (18)$$

We now consider the solution to Eq. (12) corresponding to a **wave packet**, whereby by φ and its derivatives vanish as $x\to\pm\infty$. Then, integrating Eq. (16) over the real line, and using Eqns. (17) and (18), we obtain

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\int_{-\infty}^{\infty}|\varphi(x,t)|^2\,dx &= -\frac{\hbar^2}{2m}\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial x}\left(\varphi^*\frac{\partial\varphi}{\partial x}-\varphi\frac{\partial\varphi^*}{\partial x}\right)\right]dx \\ &= -\frac{\hbar^2}{2m}\left(\varphi^*\frac{\partial\varphi}{\partial x}-\varphi\frac{\partial\varphi^*}{\partial x}\right)\Bigg|_{-\infty}^{\infty} \\ &= 0\,. \end{aligned} \qquad (19)$$

Thus, the quantity $\int_{-\infty}^{\infty}|\varphi(x,t)|^2\,dx$ is independent of time for solutions of the free-particle Schrödinger equation which correspond to wave packets.

3. (a) The multiplication of two rational numbers, m/n and p/q, where m,n,p and q are integers, yields another rational number, mp/nq, so closure is obeyed. The unit is 1, multiplication is an associative operation, and the inverse of m/n is n/m, which is a rational number. The set excludes zero, so the problem of finding the inverse of zero does not arise. Hence, the rational numbers, excluding zero, under multiplication form a group.

   (b) The sum of two non-negative integers is a non-negative integer, thus ensuring closure, addition is associative, the unit is zero, but the inverse under addition of a negative integer n is $-n$,

which is a *negative* integer and, therefore, excluded from the set. Hence, the non-negative integers under addition do not form a group.

(c) The sum of two even integers $2m$ and $2n$, where m and n are any two integers, is $2(m+n)$, which is an even integer, so closure is obeyed. Addition is associative, the unit is zero, which is an even integer, and the inverse of $2n$ is $-2n$, which is also an even integer. Hence, the even integers under addition form a group.

(d) The multiplication of two elements amounts to the addition of the integers $0, 1, \ldots, n-1$, modulo n, i.e., the addition of any two elements and, if the sum lies out side of this range, subtract n to bring it into the range. Thus, the multiplication of two nth roots of unity is again an nth root of unity, multiplication is associative, and the unit is 1. The inverse of $\mathrm{e}^{2\pi mi/n}$ is therefore

$$\mathrm{e}^{-2\pi mi/n} = \mathrm{e}^{2\pi i(n-m)/n} \tag{20}$$

Thus, the nth roots of unity form a group for any value of n.

(e) For the set of integers under ordinary subtraction, the difference between two integers n and m is another integer p, $n-m=p$, the identity is 0, since, $n-0=0$, and every integer is its own inverse, since $n-n=0$. However, subtraction is not associative because

$$(n-m)-p \neq n-(m-p) = n-m+p$$

Hence, the integers under ordinary subtraction do not form a group.

4. Let $u(x;\lambda)$ and $v(x;\lambda)$ be the fundamental solutions of the Liouville equation with the boundary conditions $y(a)=y(b)=0$. Then there are constants A and B which allow any solution y to be expressed as a linear combination of this fundamental set:

$$y(x;\lambda) = Au(x;\lambda) + Bv(x;\lambda) \tag{21}$$

These constants are determined by requiring $y(x;\lambda)$ to satisfy the boundary conditions given above. Applying these boundary conditions to the expression in (21) leads to the following equations:

$$
\begin{aligned}
y(a;\lambda) &= Au(a;\lambda) + Bv(a;\lambda) = 0 \\
y(b;\lambda) &= Au(b;\lambda) + Bv(b;\lambda) = 0
\end{aligned}
\tag{22}
$$

Equations (22) are two simultaneous equations for the unknown quantities A and B. To determine the conditions which guarantee that this system of equations has a nontrivial solution, we write these equations in matrix form:

$$
\begin{pmatrix} u(a;\lambda) & v(a;\lambda) \\ u(b;\lambda) & v(b;\lambda) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{23}
$$

Thus, we see that Equations (22) can be solved for nonzero values of A and B only if the determinant of the matrix of coefficients in (23) vanishes. Otherwise, the only solution is $A = B = 0$, which yields the trivial solution $y = 0$. The condition for a nontrivial solution of (22) is, therefore, given by

$$
\begin{vmatrix} u(a;\lambda) & v(a;\lambda) \\ u(b;\lambda) & v(b;\lambda) \end{vmatrix} = u(a;\lambda)v(b;\lambda) - u(b;\lambda)v(a;\lambda) = 0 \tag{24}
$$

This guarantees that the solution for A and B is unique. Hence, the eigenvalues are non-degenerate.

Group Theory

Problem Set 2 October 16, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1.* Show that, by requiring the existence of an identity in a group G, it is sufficient to require only a *left* identity, $ea = a$, or only a *right* identity $ae = a$, for every element a in G, since these two quantities *must* be equal.

2.* Similarly, show that it is sufficient to require only a *left* inverse, $a^{-1}a = e$, or only a *right* inverse $aa^{-1} = e$, for every element a in G, since these two quantities must also be equal.

3. Show that for any group G, $(ab)^{-1} = b^{-1}a^{-1}$.

4.* For the elements $g_1, g_2, \ldots, g_n$ of a group, determine the inverse of the n-fold product $g_1 g_2 \cdots g_n$.

5.* Show that a group is Abelian if and only if $(ab)^{-1} = a^{-1}b^{-1}$. You need to show that this condition is both necessary and sufficient for the group to be Abelian.

6. By explicit construction of multiplication tables, show that there are two distinct structures for groups of order 4. Are either of these groups Abelian?

7.* Consider the group of order 3 discussed in Section 2.4. Suppose we regard the rows of the multiplication table as individual permutations of the elements $\{e, a, b\}$ of this group. We label the permutations π_g by the group element corresponding to that row:

$$\pi_e = \begin{pmatrix} e & a & b \\ e & a & b \end{pmatrix}, \qquad \pi_a = \begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix}, \qquad \pi_b = \begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix}$$

(a) Show that, under the composition law for permutations discussed in Section 2.3, the multiplication table of the 3-element group is preserved by this association, e.g., $\pi_a \pi_b = \pi_e$.

(b) Show that for every element g in $\{e, a, b\}$,

$$\pi_g = \begin{pmatrix} e & a & b \\ g & ga & gb \end{pmatrix}$$

Hence, show that the π_g have the same multiplication table as the 3-element group.

(c) Determine the relationship between this group and S_3. This is an example of Cayley's theorem.

(d) To which of the operations on an equilateral triangle in Fig. 2.1 do these group elements correspond?

Group Theory

Solutions to Problem Set 2 October 26, 2001

1. Suppose that e is the right identity of a group G,

$$ge = g \tag{1}$$

for all g in G, and that e' is the left identity,

$$e'g = g \tag{2}$$

for all g in G. The choice $g = e'$ in the first of these equations and $g = e$ in the second, yields

$$e'e = e' \tag{3}$$

and

$$e'e = e \tag{4}$$

respectively. These equations imply that

$$e' = e \tag{5}$$

so the left and right identities are equal. Hence, we need specify only the left *or* right identity in a group in the knowledge that this is *the* identity of the group.

2. Suppose that a is the right inverse of any element g in a group G,

$$ga = e \tag{6}$$

and a' is the left inverse of g,

$$a'g = e \tag{7}$$

Multiplying the first of these equations from the left by a' and invoking the second equation yields

$$a' = a'(ga) = (a'g)a = a \tag{8}$$

so the left and right inverses of an element are equal. The same result could have been obtained by multiplying the second equation from the right by a and invoking the first equation.

3. To show that for any group G, $(ab)^{-1} = b^{-1}a^{-1}$, we begin with the properties of the inverse. We must have that

$$(ab)(ab)^{-1} = e$$

Left-multiplying both sides of this equation first by a^{-1} and then by b^{-1} yields

$$(ab)^{-1} = b^{-1}a^{-1}$$

4. For elements $g_1, g_2, \dots, g_n$ of a group G, we require the inverse of the n-fold product $g_1g_2\cdots g_n$. We proceed as in Problem 2 using the definition of the inverse to write

$$(g_1g_2\cdots g_n)(g_1g_2\cdots g_n)^{-1} = e$$

We now follow the same procedure as in Problem 2 and left-multiply both sides of this equation in turn by $g_1^{-1}, g_2^{-1}, \dots, g_n^{-1}$ to obtain

$$(g_1g_2\cdots g_n)^{-1} = g_n^{-1}\cdots g_2^{-1}g_1^{-1}$$

5. We must prove two statements here: that for an Abelian group G, $(ab)^{-1} = a^{-1}b^{-1}$, for all a and b in G, and that this equality implies that G is Abelian. If G is Abelian, then using the result

of Problem 3 and the commutativity of the composition law, we find

$$(ab)^{-1} = b^{-1}a^{-1} = a^{-1}b^{-1}$$

Now, suppose that there is a group G (which we must not *assume* is Abelian, such that

$$(ab)^{-1} = a^{-1}b^{-1}$$

for all a and b in G. We now right multiply both sides of this equation first by b and then by a to obtain

$$(ab)^{-1}ba = e$$

Then, left-multiplying both sides of this equation by (ab) yields

$$ba = ab$$

so G is Abelian. Hence, we have shown that G is an Abelian group if and only if, for elements a and b in G, $(ab)^{-1} = a^{-1}b^{-1}$.

6. To construct the multiplication table of a four-element group $\{e, a, b, c\}$ we proceed as in Section 2.4 of the course notes. The properties of the unit of the group enable us to make the following entries into the multiplication table:

	e	a	b	c
e	e	a	b	c
a	a			
b	b			
c	c			

We now consider the product aa. This cannot equal a, since that would imply that $a = e$, but it can equal any of the other elements, including the identity. However, this leads only to two *distinct* choices for the product, since the apparent difference between $aa = b$ and $aa = c$ disappears under the interchange of the

	e	a	b	c
e	e	a	b	c
a	a	e		
b	b			
c	c			

	e	a	b	c
e	e	a	b	c
a	a	b		
b	b			
c	c			

(1)

labelling of b and c. Thus, at this stage, we have two distinct structures for the multiplication table:

We now determine the remaining entries for these two groups. For the table on the left, we consider the product ab. From the Rearrangement Theorem, this cannot equal a or or e, nor can it equal b (since that would imply $a = e$). Therefore, $ab = c$, from which it follows that $ac = b$. According to the Rearrangement Theorem, the multiplication table becomes

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c		
c	c	b		

(2)

For the remaining entries of this table, we observe that $b^2 = a$ and $b^2 = e$ are equally valid assignments. These leads to two multiplication tables:

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

(3)

Note that these tables are distinct in that there is no relabelling of the elements which transforms one into the other.

We now return to the other multiplication table on the right in (1). The Rearrangement Theorem requires that the second row must be completed as follows:

	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c		
c	c	e		

(4)

Again invoking the Rearrangement Theorem, we must have that this multiplication table can be completed only as:

	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

(5)

Notice that all of the multiplication tables in (3) and (5) are Abelian and that the table in (5) is cyclic, i.e., all of the group elements can be obtained by taking successive products of any non-unit element.

We now appear to have three distinct multiplication tables for groups of order 4: the two tables in (3) and the one in (5). However, if we reorder the elements in the second table in (3) to $\{e, b, a, c\}$ and reassemble the multiplication table (using the same products), we obtain

	e	b	a	c
e	e	b	a	c
c	b	a	c	e
b	a	c	e	b
a	c	e	b	a

(6)

which, under the relabelling $a \mapsto b$ and $b \mapsto a$, is identical to (5). Hence, there are only two distinct structures of groups with four elements:

$$\begin{array}{c|cccc} & e & a & b & c \\ \hline e & e & a & b & c \\ a & a & e & c & b \\ b & b & c & e & a \\ c & c & b & a & e \end{array} \qquad \begin{array}{c|cccc} & e & a & b & c \\ \hline e & e & a & b & c \\ a & a & b & c & e \\ b & b & c & e & a \\ c & c & e & a & b \end{array} \tag{7}$$

7. (a) All of the products involving the identity are self-evident. The only products that must be calculated explicitly are a^2, ab, ba, and b^2. These are given by

$$\begin{aligned} \pi_a\pi_a &= \begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix}\begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix} = \begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix} = \pi_b \\ \pi_a\pi_b &= \begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix}\begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix} = \begin{pmatrix} e & a & b \\ e & a & b \end{pmatrix} = \pi_e \\ \pi_b\pi_a &= \begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix}\begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix} = \begin{pmatrix} e & a & b \\ e & a & b \end{pmatrix} = \pi_e \\ \pi_b\pi_b &= \begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix}\begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix} = \begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix} = \pi_a \end{aligned}$$

Thus, the association $g \to \pi_g$, for $g = e, a, b$ preserves the products of the three-element group.

(b) With the construction

$$\pi_g = \begin{pmatrix} e & a & b \\ g & ga & gb \end{pmatrix} \tag{9}$$

for $g = e, a, b$, we can use the rows of the multiplication table on p. 19 to obtain

$$\begin{aligned} \pi_e &= \begin{pmatrix} e & a & b \\ ee & ea & eb \end{pmatrix} = \begin{pmatrix} e & a & b \\ e & a & b \end{pmatrix} \\ \pi_a &= \begin{pmatrix} e & a & b \\ ae & aa & ab \end{pmatrix} = \begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix} \end{aligned}$$

$$\pi_b = \begin{pmatrix} e & a & b \\ be & ba & bb \end{pmatrix} = \begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix} \tag{10}$$

Thus, the association $g \rightarrow \pi_g$ is one-to-one and preserves the products of the 3-element group. Hence, these groups are equivalent.

(c) The elements of the three-element group correspond to the **cyclic permutations** of S_3. In other words, given a reference order $\{a, b, c\}$, the cyclic permutations are $a \mapsto b$, $b \mapsto c$, and $c \mapsto a$, yielding $\{b, c, a\}$, and then $b \mapsto c$, $c \mapsto a$, and $a \mapsto b$, yielding $\{c, a, b\}$.

(d) These elements correspond to the **rotations** of an equilateral triangle, i.e., the elements $\{e, d, f\}$ in Fig. 2.1.

Group Theory

Problem Set 3 October 23, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1.* List all of the subgroups of any group whose order is a prime number.

2.* Show that a group whose order is a prime number is necessarily cyclic, i.e., all of the elements can be generated from the powers of any non-unit element.

3. Suppose that, for a group G, $|G| = pq$, where p and q are both prime. Show that every proper subgroup of G is cyclic.

4.* Let g be an element of a finite group G. Show that $g^{|G|} = e$.

5. In a quotient group G/H, which set *always* corresponds to the unit "element"?

6. Show that, for an Abelian group, every element is in a class by itself.

7. Show that every subgroup with index 2 is self-conjugate.

Hint: The conjugating element is either in the subgroup or not. Consider the two cases separately.

8.* Consider the following cyclic group of order 4, $G = \{a, a^2, a^3, a^4 = e\}$ (cf. Problem 6, Problem Set 2). Show, by direct multiplication or otherwise, that the subgroup $H = \{e, a^2\}$ is self-conjugate and identify the elements in the factor group G/H.

9.* Suppose that there is an isomorphism ϕ from a group G onto a group G'. Show that the identity e of G is mapped onto the identity e' of G': $e' = \phi(e)$.

Hint: Use the fact that $e = ee$ must be preserved by ϕ and that $\phi(g) = e'\phi(g)$ for all g in G.

Group Theory

Solutions to Problem Set 3 November 2, 2001

1. According to Lagrange's theorem, the order of a subgroup H of a group G must be a divisor of $|G|$. Since the divisors of a prime number are only the number itself and unity, the subgroups of a group of prime order must be either the unit element alone, $H = \{e\}$, or the group G itself, $H = G$, both of which are *improper* subgroups. Therefore, a group of prime order has no *proper* subgroups.

2. From a group G of prime order, select any element g, which is not the unit element, and form its period:
$$g, g^2, g^3, \ldots, g^n = e\,,$$
where n is the order of g (Sec. 2.4). The period *must* include every element in G, because otherwise we would have constructed a subgroup whose order is neither unity nor $|G|$. This contradicts the conclusion of Problem 1. Hence, a group of prime order is necessarily cyclic (but a cyclic group need not necessarily be of prime order).

3. For a group G with $|G| = pq$, where p and q are both prime, we know from Lagrange's theorem that the only proper subgroups have order p and q. Since these subgroups are of prime order, the conclusion of Problem 2 requires these subgroups to be cyclic.

4. Since the period of an element g of a group G forms a subgroup of G (this is straightforward to verify), Lagrange's theorem requires that $|g|$ must be a divisor of $|G|$, i.e., $G = k|g|$ for some integer k. Hence,
$$g^{|G|} = g^{k|g|} = (g^{|g|})^k = e^k = e\,.$$

5. The identity e of a group G has the property that for every element g in G, $ag = ge = g$. We also have that different cosets either have no common elements or have only common elements. Thus, in the factor group G/H of G generated by a subgroup H, the set which contains the unit element corresponds to the unit element of the factor group, since

$$\{e, h_1, h_2, \ldots\}\{a, b, c, \ldots\} = \{a, b, c, \ldots\}\,.$$

6. The class of an element a in a group G is defined as the set of elements gag^{-1} for all elements g in G. If G is Abelian, then we have

$$gag^{-1} = gg^{-1}a = a$$

for all g in G. Hence, in an Abelian group, every element is in a class by itself.

7. Let H be a subgroup of a group of G of index 2, i.e., H has two left cosets and two right cosets. If H is self-conjugate, then $gHg^{-1} = H$ for any g in G. Therefore, to show that H is self-conjugate, we must show that $gH = Hg$ for any g in G, i.e., that the left and right cosets are the same. Since H has index 2, and H is itself a right coset, all of the elements in Hg must either be in H or in the other coset of H, which we will call A. There two possibilities: either g is in H or g is not in H. If g is an element of H, then, according to the Rearrangement Theorem,

$$Hg = gH = H\,.$$

If g is not in H, then it is in A, which is a right coset of H. Two (left or right) cosets of a subgroup have either all elements in common or no elements in common. Thus, since the unit element

must be contained in H, the set Hg will contain g which, by hypothesis, is in A. We conclude that

$$Hg = gH = A\,.$$

Therefore,

$$Hg = gH$$

for all g in G and H is, therefore, a self-conjugate subgroup.

8. The subgroup $H = \{e, a^2\}$ of the group $G = \{e, a, a^2, a^3, a^4 = e\}$ has index 2. Therefore, according to Problem 7, H *must* be self-conjugate. Therefore, the elements of the factor group G/H are the subgroup H, which corresponds to the unit element, so we call it $\mathcal{E}$, and the set consisting of the elements $\mathcal{A} = \{a, a^3\}$: $G/H = \{\mathcal{E}, \mathcal{A}\}$.

9. Let ϕ be an isomorphism between a group G and a group G', i.e. ϕ is a one-to-one mapping between all the elements g of G and g' of G'. From the group properties we have that the identity e of G must obey the relation

$$e = ee\,.$$

Since ϕ preserves all products, this relation must in particular be preserved by ϕ:

$$\phi(e) = \phi(e)\phi(e)\,.$$

The group properties require that, for any element g of G,

$$\phi(g) = e'\phi(g)\,,$$

where e' is the identity of G'. Setting $g = e$ and comparing with the preceding equation yields the equality

$$\phi(e)\phi(e) = e'\phi(e)\,,$$

which, by cancellation, implies

$$e' = \phi(e)\,.$$

Thus, an isomorphism maps the identity in G onto the identity in G'.

Group Theory

Problem Set 4 October 30, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1.* Given a set of matrices $D(g)$ that form a representation a group G, show that the matrices which are obtainable by a similarity transformation $UD(g)U^{-1}$ are also a representation of G.

2.* Show that the trace of three matrices A, B, and C satisfies the following relation:

$$\mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)$$

3. Generalize the result in Problem 4 to show that the trace of an n-fold product of matrices is invariant under cyclic permutations of the product.

4.* Show that the trace of an arbitrary matrix A is invariant under a similarity transformation UAU^{-1}.

5. Consider the following representation of S_3:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad a = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad b = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$$

$$c = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad d = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}, \quad f = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$$

How can these matrices be permuted to provide an equally faithful representation of S_3? Relate your result to the class identified with each element.

6.* Consider the planar symmetry operations of an equilateral triangle. Using the matrices in Example 3.2 determined from transformations of the coordinates in Fig. 3.1, construct a *three*-dimensional representation of S_3 in the (x, y, z) coordinate system, where the z-axis emanates from the geometric center of the triangle. Is this representation reducible or irreducible? If it is reducible determine the irreducible representations which form the direct sum of this representation.

7. Show that two matrices are simultaneously diagonalizable if and only if they commute.

Hint: Two matrices A and B are simultaneously diagonalizable if the *same* similarity transformation brings both matrices into a form where they have only diagonal entries. Proving that simultaneous diagonalizability implies commutativity is straightforward. To prove the converse, suppose that there is a similarity transformation which brings *one* of the matrices into diagonal form. By writing out the matrix elements of the products and using the fact that A and B commute, show that the same similarity transformation *must* also diagonalize the other matrix.

8.* What does the result of Problem 7 imply about the dimensionalities of the irreducible representations of Abelian groups?

9.* Verify that the matrices

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$$

form a representation for the two-element group $\{e, a\}$. Is this representation reducible or irreducible? If it is reducible determine the one-dimensional representations which form the direct sum of this representation.

10.* Prove the relations in Eqns (3.9) and (3.11).

Group Theory

Solutions to Problem Set 4 November 9, 2001

1. Let $D'(g) = UD(g)U^{-1}$, where $D(g)$ is a representation of a group G with elements g. To show that $D'(g)$ is also a representation of G, it is sufficient to show that this representation preserves the multiplication table of G. Thus, let a and b be any two elements of G with matrix representations $D(a)$ and $D(b)$. The product ab is represented by

$$D(ab) = D(a)D(b)\,.$$

Therefore,

$$\begin{aligned} D'(ab) &= UD(ab)U^{-1} \\ &= UD(a)D(b)U^{-1} \\ &= UD(a)U^{-1}UD(b)U' \\ &= D'(a)D'(b)\,, \end{aligned}$$

so multiplication is preserved and $D'(g)$ is therefore also a representation of G.

2. The trace of a matrix is the sum of its diagonal elements. Therefore, the trace of the product of three matrices A, B, and C is given by

$$\mathrm{tr}(ABC) = \sum_{ijk} A_{ij}B_{jk}C_{ki}\,.$$

By using the fact that i, j, and k are dummy summation indices with the same range, this sum can be written in the equivalent forms

$$\sum_{ijk} A_{ij}B_{jk}C_{ki} = \sum_{ijk} C_{ki}A_{ij}B_{jk} = \sum_{ijk} B_{jk}C_{ki}A_{ij}\,.$$

But the second and third of these are

$$\sum_{ijk} C_{ki} A_{ij} B_{jk} = \mathrm{tr}(CAB)$$

and

$$\sum_{ijk} B_{jk} C_{ki} A_{ij} = \mathrm{tr}(BCA)\,,$$

respectively. Thus, we obtain the relation

$$\mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)\,.$$

3. The trace of an n-fold product, $A_1 A_2 \cdots A_n$ is

$$\mathrm{tr}(A_1 A_2 \cdots A_n) = \sum_{i_1, i_2, \ldots i_n} (A_1)_{i_1 i_2} (A_2)_{i_2 i_3} \cdots (A_n)_{i_n i_1}\,.$$

Proceeding as in Problem 2, we observe that the i_k $(k = 1, \ldots, n)$ are dummy summation indices all of which have the same range. Thus, any cyclic permutation of the matrices in the product leaves the sum and, hence, the trace invariant.

4. From Problem 2, we have that

$$\mathrm{tr}(UAU^{-1}) = \mathrm{tr}(U^{-1}UA) = \mathrm{tr}(A)\,,$$

so a similarity transformation leaves the trace of a matrix invariant.

5. Given a faithful representation of a group, similarity transformations of the matrices provide equally faithful representations. Since we wish to obtain permutations of a particular matrix representation, we base our similarity transformations on the non-unit elements in the group. Thus, consider the following similarity transformations:

$$\begin{aligned}
a\{e,a,b,c,d,f\}a^{-1} &= \{e,a,c,b,f,d\}\,, \\
b\{e,a,b,c,d,f\}b^{-1} &= \{e,c,b,a,f,d\}\,, \\
c\{e,a,b,c,d,f\}c^{-1} &= \{e,b,a,c,f,d\}\,, \\
d\{e,a,b,c,d,f\}d^{-1} &= \{e,b,c,a,d,f\}\,, \\
e\{e,a,b,c,d,f\}e^{-1} &= \{e,c,a,b,d,f\}\,.
\end{aligned} \tag{1}$$

Thus, the following permutations of the elements $\{e,a,b,c,d,f\}$ provide equally faithful representations:

$$\begin{aligned}
&\{e,a,c,b,f,d\}\,, \quad \{e,c,b,a,f,d\}\,, \\
&\{e,b,a,c,f,d\}\,, \quad \{e,b,c,a,d,f\}\,, \\
&\{e,c,a,b,d,f\}\,.
\end{aligned}$$

Notice that only elements within the same class can be permuted. For S_3, the classes are $\{e\}$, $\{a,b,c\}$, $\{d,f\}$.

6. In the basis (x,y,z) where x and y are given in Fig. 3.1 and the z-axis emanates from the origin of this coordinate system (the geometric center of the triangle), all of the symmetry operations of the equilateral triangle leave the z-axis invariant. This is because the z-axis is either an axis of rotation (for operations d and f) or lies within the reflection plane (for operations a, b, and c). Hence,

the matrices of these operations are given by

$$e = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad a = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} & 0 \\ -\sqrt{3} & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

$$b = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} & 0 \\ \sqrt{3} & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad c = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$d = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} & 0 \\ \sqrt{3} & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad f = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} & 0 \\ -\sqrt{3} & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

This representation is seen to be *reducible* and that it is the direct sum of the representation in Example 3.2 (which, as discussed in Example 3.4, is irreducible) and the identical representation.

7. We must show (i) that two matrices which are simultaneously diagonalizable commute and (ii) that two matrices which commute are simultaneously diagonalizable. Showing (i) is straightforward. For two $d \times d$ matrices A and B which are simultaneously diagonalizable, there is a matrix U such that

$$UAU^{-1} = D_A \qquad \text{and} \qquad UBU^{-1} = D_B\,,$$

where D_A and D_B are diagonal forms of these matrices. Clearly, therefore, we have that

$$D_A D_B = D_B D_A\,.$$

Hence, transforming back to the original basis,

$$\underbrace{(U^{-1}D_AU)}_{A}\,\underbrace{(U^{-1}D_BU)}_{B} = \underbrace{(U^{-1}D_BU)}_{B}\,\underbrace{(U^{-1}D_AU)}_{A},$$

so A and B commute.

Now suppose that A and B commute and there is a transformation that brings *one* of these matrices, say A, into the diagonal form D_A:

$$UAU^{-1} = D_A\,.$$

Then, with

$$UBU^{-1} = B'\,,$$

the commutation relation $AB = BA$ transforms to

$$D_A B' = B' D_A\,.$$

The (i,j)th matrix element of these products is

$$\begin{aligned}(D_A B')_{ij} &= \sum_k (D_A)_{ik}(B')_{kj} = (D_A)_{ii}(B')_{ij} \\ = (B'D_A)_{ij} &= \sum_k (B')_{ik}(D_A)_{kj} = (B')_{ij}(D_A)_{jj}\,.\end{aligned}$$

After a simple rearrangement, we have

$$(B')_{ij}[(D_A)_{ii} - (D_A)_{jj}] = 0\,.$$

There are three cases to consider:

Case I. All of the diagonal entries of D_A are distinct. Then,

$$(D_A)_{ii} - (D_A)_{jj} \neq 0 \qquad \text{if } i \neq j\,,$$

so *all* of the off-diagonal matrix elements of B' vanish, i.e., B' is a diagonal matrix. Thus, the same similarity transformation which diagonalizes A also diagonalizes B.

Case II. All of the diagonal entries of D_A are the same. In this case D_A is proportional to the unit matrix, $D_A = cI$, for some complex constant c. Hence, this matrix is *always* diagonal,

$$U(cI)U^{-1} = cI$$

and, in particular, it is diagonal when B is diagonal.

Case III. Some of the diagonal entries are the same and some are distinct. If we arrange the elements of D_A such that the first p elements are the same, $(D_A)_{11} = (D_A)_{22} = \cdots = (D_A)_{pp}$, then D_A has the general form

$$D_A = \begin{pmatrix} cI_p & 0 \\ 0 & D'_A \end{pmatrix},$$

where I_p is the $p \times p$ unit matrix and c is a complex constant. From Cases I and II, we deduce that B must be of the form

$$B = \begin{pmatrix} B_p & 0 \\ 0 & D'_B \end{pmatrix},$$

where B_p is some $p \times p$ matrix and D'_B is a diagonal matrix. Let V_p be the matrix which diagonalizes B:

$$V_p B_p V_p^{-1} = D''_B\,.$$

Then the matrix

$$V = \begin{pmatrix} V_p & 0 \\ 0 & I_{d-p} \end{pmatrix}$$

diagonalizes B while leaving D_A unchanged. Here, I_{d-p} is the $(d-p) \times (d-p)$ unit matrix.

Hence, in all three cases, we have shown that the same transformation which diagonalizes A also diagonalizes B.

8. The matrices of any representation $\{A_1, A_2, \ldots, A_n\}$ of an Abelian group G commute:

$$A_i A_j = A_j A_i$$

for all i and j. Hence, according to Problem 7, these matrices can all be simultaneously diagonalized. Since this is true of *all* representations of G, we conclude that all irreducible representations of Abelian groups are one-dimensional, i.e., they are numbers with ordinary multiplication as the composition law.

9. To verify that the matrices

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} \tag{2}$$

form a representation for the two-element group $\{e, a\}$, we need to demonstrate that the multiplication table for this group,

	e	a
e	e	a
a	a	e

is fulfilled by these matrices. The products $e^2 = e$, $ea = a$, and $ae = a$ can be verified by inspection. The product a^2 is

$$a^2 = \tfrac{1}{4}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = e\,,$$

so the matrices in (2) form a representation of the two-element group.

Since these matrices commute, they can be diagonalized simultaneously (Problem 7). Since the matrix is the unit matrix, we can determine the diagonal form of a, simply by finding its eigenvalues. The characteristic equation of a is

$$\begin{aligned} det(a - \lambda I) &= \begin{vmatrix} -\frac{1}{2} - \lambda & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} - \lambda \end{vmatrix} \\ &= -(\tfrac{1}{2} - \lambda)(\tfrac{1}{2} + \lambda) - \tfrac{3}{4} = \lambda^2 - 1\,. \end{aligned}$$

which yields $\lambda = \pm 1$. Therefore, diagonal form of a is

$$a = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3}$$

so this representation is the direct sum of the identical representation $\{1, 1\}$, and the "parity" representation $\{1, -1\}$. Note that, according to Problem 8, *every* representation of the two-element group with dimensionality greater than two *must* be reducible.

10. The relations in (3.9) and (3.11) can be proven simultaneously, since they differ only by complex conjugation, which preserves the order of matrices. The (i, j)th matrix element of n-fold product of matrices $A_1, A_2, \ldots, A_n$ is

$$(A_1 A_2 \cdots A_n)_{ij} = \sum_{k_1, k_2, \ldots, k_{n-1}} (A_1)_{ik_1} (A_2)_{k_1 k_2} \cdots (A_n)_{k_{n-1} j}\,.$$

The corresponding matrix element of the transpose of this product is

$$[(A_1 A_2 \cdots A_n)^{\mathrm{t}}]_{ij} = (A_1 A_2 \cdots A_n)_{ji}\,.$$

Thus, since the k_i are dummy indices,

$$\begin{aligned} [(A_1 A_2 \cdots A_n)^{\mathrm{t}}]_{ij} &= \sum_{k_1, k_2, \ldots, k_{n-1}} (A_1)_{jk_1} (A_2)_{k_1 k_2} \cdots (A_n)_{k_{n-1} i} \\ &= \sum_{k_1, k_2, \ldots, k_{n-1}} (A_1^{\mathrm{t}})_{k_1 j} (A_2^{\mathrm{t}})_{k_2 k_1} \cdots (A_n^{\mathrm{t}})_{i k_{n-1}} \\ &= \sum_{k_1, k_2, \ldots, k_{n-1}} (A_n^{\mathrm{t}})_{i k_{n-1}} (A_{n-1}^{\mathrm{t}})_{k_{n-1} k_{n-2}} \cdots (A_2^{\mathrm{t}})_{k_2 k_1} (A_1^{\mathrm{t}})_{k_1 j} \end{aligned}$$

We conclude that

$$(A_1 A_2 \cdots A_n)^{\mathrm{t}} = A_n^{\mathrm{t}} A_{n-1}^{\mathrm{t}} \cdots A_1^{\mathrm{t}}$$

and, similarly, that

$$(A_1 A_2 \cdots A_n)^\dagger = A_n^\dagger A_{n-1}^\dagger \cdots A_1^\dagger$$

Group Theory

Problem Set 5 November 6, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1. In proving Theorem 3.2, we established the relation $B_iB_i^\dagger = I$. Using the definitions in that proof, show that this result implies that $B_i^\dagger B_i = I$ as well.

Hint: Show that $B_iB_i^\dagger = I$ implies that $\tilde{A}_i D \tilde{A}_i^\dagger = D$.

2.* Consider the three-element group $G = \{e, a, b\}$ (Sec. 2.4).

(a) Show that this group is Abelian and cyclic (cf. Problem 2, Problem Set 3).

(b) Consider a one-dimensional representation based on choosing $a = z$, where z is a complex number. Show that for this to produce a representation of G, we must require that $z^3 = 1$.

(c) Use the result of (b) to obtain three representations of G. Given what you know about the irreducible representations of Abelian groups (Problem 8, Problem Set 4), are there any other irreducible representations of G?

3.* Generalize the result of Problem 2 to any cylic group of order n.

4.* Use Schur's First Lemma to prove that all the irreducible representations of an Abelian group are one-dimensional.

5.* Consider the following matrices:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}, \quad b = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}$$

$$c = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}, \qquad d = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Verify that these matrices form a representation of S_3. Use Schur's first Lemma to determine if this representation reducible or irreducible. If reducible, determine the irreducible representations that are obtained from the diagonal form of these matrices.

Group Theory

Solutions to Problem Set 5 November 16, 2001

1. In proving Theorem 3.2, we established that $B_iB_i^\dagger = I$, where

$$B_i = D^{-1/2}\tilde{A}_iD^{1/2}$$

To show that this result implies that $B_i^\dagger B_i = I$, we first use the definitions of B_i and $B_i^\dagger$ to write

$$B_i^\dagger B_i = D^{1/2}\tilde{A}_i^\dagger D^{-1}\tilde{A}_iD^{1/2}$$

We can find an expression for D^{-1} by first rearranging

$$B_iB_i^\dagger = D^{-1/2}\tilde{A}_iD\tilde{A}_i^\dagger D^{-1/2} = I$$

as

$$\tilde{A}_iD\tilde{A}_i^\dagger = D$$

Then, taking the inverse of both sides of this equation yields

$$D^{-1} = \tilde{A}_i^{\dagger -1}D^{-1}\tilde{A}_i^{-1}$$

Therefore,

$$\begin{aligned} B_i^\dagger B_i &= D^{1/2}\tilde{A}_i^\dagger D^{-1}\tilde{A}_iD^{1/2} \\ &= D^{1/2}\tilde{A}_i^\dagger(\tilde{A}_i^{\dagger -1}D^{-1}\tilde{A}_i^{-1})\tilde{A}_iD^{1/2} \\ &= I \end{aligned}$$

2. (a) The multiplication table for the three-element group is shown below (Sec. 2.4):

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

We can see immediately that $a^2 = b$ and that $ab = ba = a^3 = e$, Thus, the three-element group can be written as $\{a, a^2, a^3 = e\}$, i.e., it is a cyclic group (and, therefore, Abelian).

(b) By choosing a one-dimensional representation $a = z$, for some complex number z, the multiplication table requires that $a^3 = e$, which means that $z^3 = 1$.

(c) There are three solutions to $z^3 = 1$: $z = 1, \mathrm{e}^{2\pi i/3}, and \mathrm{e}^{4\pi i/3}$. The three irreducible representations are obtained by choosing $a = 1$, $a = \mathrm{e}^{2\pi i/3}$, and $a = \mathrm{e}^{4\pi i/3}$. Denoting these representations by Γ_1, Γ_2, and Γ_3, we obtain

	e	a	b
Γ_1	1	1	1
Γ_2	1	$\mathrm{e}^{2\pi i/3}$	$\mathrm{e}^{4\pi i/3}$
Γ_2	1	$\mathrm{e}^{4\pi i/3}$	$\mathrm{e}^{2\pi i/3}$

3. The preceding Problem can be generalized to any cyclic group of order n. The elements of this group are $\{g, g^2, \ldots, g^n = e\}$. By writing $g = z$, we require that $z^n = 1$. The solutions to this equation are the nth roots of unity:

$$z = \mathrm{e}^{2m\pi i/n}, \quad m = 0, 1, 2, \ldots, n-1$$

Accordingly, there are n irreducible representations based on the n choices $z = \mathrm{e}^{2m\pi i/n}$ together with the requirements of the group multiplication table.

4. Suppose that we have a representation of an Abelian group of dimensionality d is greater than one. Suppose furthermore that these matrices are not all unit matrices (for, if they were, the representation would already be reducible to the d-fold direct sum of the identical representation.) Then, since the group is Abelian, and the representation must reflect this fact, *any* non-unit matrix in the representation commutes with all the other matrices in the representation. According to Schur's First Lemma, this renders the representation reducible. Hence, all the irreducible representations of an Abelian group are one-dimensional.

5. For the following matrices,

$$e = d = f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = b = c = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}$$

to be a representation of S_3, their products must preserve the multiplication table of this group, which was discussed in Sec. 2.4 and is displayed below:

	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	e	d	f	b	c
b	b	f	e	d	c	a
c	c	d	f	e	a	b
d	d	c	a	b	f	e
f	f	b	c	a	e	d

To determine the multiplication table of this representation, we use the notation E for the unit matrix, corresponding to the elements e, d, and f, and A for the matrix corresponding to the elements a, b, and c. Then, by observing that $A^2 = E$ (as is required by the group multiplication table) the multiplication table of this representation is straightforward to calculate, and is shown below:

	e	a	b	c	d	f
e	E	A	A	A	E	E
a	A	E	E	E	A	A
b	A	E	E	E	A	A
c	A	E	E	E	A	A
d	E	A	A	A	E	E
f	E	A	A	A	E	E

If we now take the multiplication table of S_3 and perform the mapping $\{e, d, f\} \mapsto E$ and $\{a, b, c\} \mapsto A$, we get the same table as that just obtained by calculating the matrix products directly. Hence, these matrices form a representation of S_3.

From Schur's First Lemma, we see that the matrix

$$\tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}$$

commutes with all the matrices of the representation. Since this is not a unit matrix, the representation must be reducible.

The diagonal form of these matrices must have entries of one-dimensional irreducible representations. Two one-dimensional irreducible representations of S_3 (we will see later that these are the *only* irreducible representations of S_3) are (Example 3.2) the identical representation,

$$A_e = 1, \quad A_a = 1, \quad A_b = 1,$$

$$A_c = 1, \quad A_d = 1, \quad A_f = 1$$

and the 'parity' representation,

$$A_e = 1, \quad A_a = -1, \quad A_b = -1,$$

$$A_c = -1, \quad A_d = 1, \quad A_f = 1$$

Since the diagonal forms of the matrices are obtained by performing a similarity transformation on the original matrices, which

preserves the trace, they must take the form

$$e = d = f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = b = c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

i.e., this reducible representation contains both the identical and parity representations.

Group Theory

Problem Set 6 November 13, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1.* Verify the Great Orthogonality Theorem for the following irreducible representation of S_3:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = \tfrac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \qquad b = \tfrac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$$

$$c = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad d = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \qquad f = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$$

2.* Does the following representation of the three-element group $\{e, a, b\}$:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad a = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad b = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$$

satisfy the Great Orthogonality Theorem? Explain your answer.

3.* Specialize the Great Orthogonality Theorem to Abelian groups. When viewed as the components of a vector in a $|G|$-dimensional space, what does the Great Orthogonality Theorem state about the relationship between different irreducible representations? What bound does this place on the number of irreducible representations of an Abelian group?

4.* Consider the irreducible representations of the three-element calculated in Problem 2 of Problem Set 5.

(a) Verify that the Great Orthogonality Theorem, in the reduced form obtained in Problem 3, is satisfied for these representations.

(b) In view of the discussion in Sec. 4.4, would you expect to find any other irreducible representations of this group?

(c) Would you expect your answer in (b) to apply to cyclic groups of any order?

5.* Consider any Abelian group. By using the notion of the order of an element (Sec. 2.4), determine the *magnitude* of every element in a representation. Is this consistent with the Great Orthogonality Theorem?

Group Theory

Solutions to Problem Set 6 November 23, 2001

1. The Great Orthogonality Theorem states that, for the matrix elements of the same irreducible representation $\{A_1, A_2, \ldots, A_{|G|}\}$ of a group G,

$$\sum_\alpha (A_\alpha)_{ij}(A_\alpha)^*_{i'j'} = \frac{|G|}{d}\delta_{i,i'}\delta_{j,j'}\,.$$

Thus, we first form the vectors $\boldsymbol{V}_{ij}$ whose components are the (i,j)th elements taken from each matrix in the representation in some fixed order. The Great Orthogonality Theorem can then be expressed more concisely as

$$\boldsymbol{V}_{ij}\cdot\boldsymbol{V}^*_{i'j'} = \frac{|G|}{d}\delta_{i,i'}\delta_{j,j'}\,.$$

For the given representation of S_3,

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a = \tfrac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad b = \tfrac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix},$$

$$c = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad d = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \qquad f = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix},$$

these vectors are:

$$\boldsymbol{V}_{11} = \left(1, \tfrac{1}{2}, \tfrac{1}{2}, -1, -\tfrac{1}{2}, -\tfrac{1}{2}\right),$$

$$\boldsymbol{V}_{12} = \left(0, -\tfrac{1}{2}\sqrt{3}, \tfrac{1}{2}\sqrt{3}, 0, \tfrac{1}{2}\sqrt{3}, -\tfrac{1}{2}\sqrt{3}\right),$$

$$\boldsymbol{V}_{21} = \left(0, -\tfrac{1}{2}\sqrt{3}, \tfrac{1}{2}\sqrt{3}, 0, -\tfrac{1}{2}\sqrt{3}, \tfrac{1}{2}\sqrt{3}\right),$$

$$\boldsymbol{V}_{22} = \left(1, -\tfrac{1}{2}, -\tfrac{1}{2}, 1, -\tfrac{1}{2}, -\tfrac{1}{2}\right).$$

Note that, since all of the entries are real, complex conjugation is not required for substitution into the Great Orthogonality Theorem. For $i = j$ and $i' = j'$, with $|G| = 6$ and $d = 2$, we have

$$\begin{aligned}
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{11} &= 1 + \tfrac{1}{4} + \tfrac{1}{4} + 1 + \tfrac{1}{4} + \tfrac{1}{4} = 3\,,\\
\boldsymbol{V}_{12} \cdot \boldsymbol{V}_{12} &= 0 + \tfrac{3}{4} + \tfrac{3}{4} + 0 + \tfrac{3}{4} + \tfrac{3}{4} = 3\,,\\
\boldsymbol{V}_{21} \cdot \boldsymbol{V}_{21} &= 0 + \tfrac{3}{4} + \tfrac{3}{4} + 0 + \tfrac{3}{4} + \tfrac{3}{4} = 3\,.\\
\boldsymbol{V}_{22} \cdot \boldsymbol{V}_{22} &= 1 + \tfrac{1}{4} + \tfrac{1}{4} + 1 + \tfrac{1}{4} + \tfrac{1}{4} = 3\,,
\end{aligned}$$

all of which are in accord with the Great Orthogonality Theorem. For $i \neq j$ and/or $i' \neq j'$, we have

$$\begin{aligned}
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{12} &= 0 - \tfrac{1}{4}\sqrt{3} + \tfrac{1}{4}\sqrt{3} + 0 - \tfrac{1}{4}\sqrt{3} + \tfrac{1}{4}\sqrt{3} = 0\,,\\
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{21} &= 0 - \tfrac{1}{4}\sqrt{3} + \tfrac{1}{4}\sqrt{3} + 0 + \tfrac{1}{4}\sqrt{3} - \tfrac{1}{4}\sqrt{3} = 0\,,\\
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{22} &= 1 - \tfrac{1}{4} - \tfrac{1}{4} - 1 + \tfrac{1}{4} + \tfrac{1}{4} = 0\,,\\
\boldsymbol{V}_{12} \cdot \boldsymbol{V}_{21} &= 0 + \tfrac{3}{4} + \tfrac{3}{4} + 0 - \tfrac{3}{4} - \tfrac{3}{4} = 0\,,\\
\boldsymbol{V}_{12} \cdot \boldsymbol{V}_{22} &= 0 + \tfrac{1}{4}\sqrt{3} - \tfrac{1}{4}\sqrt{3} + 0 - \tfrac{1}{4}\sqrt{3} + \tfrac{1}{4}\sqrt{3} = 0\,,\\
\boldsymbol{V}_{21} \cdot \boldsymbol{V}_{22} &= 0 + \tfrac{1}{4}\sqrt{3} - \tfrac{1}{4}\sqrt{3} + 0 + \tfrac{1}{4}\sqrt{3} - \tfrac{1}{4}\sqrt{3} = 0\,,
\end{aligned}$$

which is also in accord with the Great Orthogonality Theorem.

2. For the following two-dimensional representation of the three-element group,

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad a = \tfrac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad b = \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix},$$

we again form the vectors $\boldsymbol{V}_{ij}$ whose components are the (i,j)th elements of each matrix in the representation:

$$\boldsymbol{V}_{11} = \left(1, -\tfrac{1}{2}, -\tfrac{1}{2}\right),$$

$$\begin{aligned}
\boldsymbol{V}_{12} &= \left(0, \tfrac{1}{2}\sqrt{3}, -\tfrac{1}{2}\sqrt{3}\right), \\
\boldsymbol{V}_{21} &= \left(0, -\tfrac{1}{2}\sqrt{3}, \tfrac{1}{2}\sqrt{3}\right), \\
\boldsymbol{V}_{22} &= \left(1, -\tfrac{1}{2}, -\tfrac{1}{2}\right).
\end{aligned}$$

Calculating the summation in the Great Orthogonality Theorem, first with $i = j$ and $i' = j'$, we have

$$\begin{aligned}
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{11} &= 1 + \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{3}{2}, \\
\boldsymbol{V}_{12} \cdot \boldsymbol{V}_{12} &= 0 + \tfrac{3}{4} + \tfrac{3}{4} = \tfrac{3}{2}, \\
\boldsymbol{V}_{21} \cdot \boldsymbol{V}_{12} &= 0 + \tfrac{3}{4} + \tfrac{3}{4} = \tfrac{3}{2}, \\
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{11} &= 1 + \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{3}{2},
\end{aligned}$$

all of which are in accord with the Great Orthogonality Theorem with $|G| = 3$ and $d = 2$. Performing the analogous summations with $i \neq j$ and/or $i' \neq j'$, yields

$$\begin{aligned}
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{12} &= 0 - \tfrac{1}{4}\sqrt{3} + \tfrac{1}{4}\sqrt{3} = 0, \\
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{21} &= 0 + \tfrac{1}{4}\sqrt{3} - \tfrac{1}{4}\sqrt{3} = 0, \\
\boldsymbol{V}_{11} \cdot \boldsymbol{V}_{22} &= 1 + \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{3}{2}, \\
\boldsymbol{V}_{12} \cdot \boldsymbol{V}_{21} &= 0 - \tfrac{3}{4} - \tfrac{3}{4} = -\tfrac{3}{2}, \\
\boldsymbol{V}_{12} \cdot \boldsymbol{V}_{22} &= 0 - \tfrac{1}{4}\sqrt{3} + \tfrac{1}{4}\sqrt{3} = 0, \\
\boldsymbol{V}_{21} \cdot \boldsymbol{V}_{22} &= 0 + \tfrac{1}{4}\sqrt{3} - \tfrac{1}{4}\sqrt{3} = 0,
\end{aligned}$$

which is *not* consistent with the Great Orthogonality Theorem, since *all* of these quantities must vanish. If there is even a single violation of the Great Orthogonality Theorem, as is the case here, the representation is necessarily reducible.

3. All of the irreducible representations of an Abelian group are one-dimensional (e.g., Problem 4, Problem Set 5). Hence, for Abelian groups, the Great Orthogonality Theorem reduces to

$$\sum_\alpha A_\alpha^k A_\alpha^{k'*} = |G|\delta_{k,k'} \,.$$

If we view the irreducible representations as $|G|$-dimensional vectors $\boldsymbol{A}^k$ with entries A_α^k,

$$\boldsymbol{A}^k = (A_1^k, A_2^k, \ldots A_{|G|}^k) \,,$$

then the Great Orthogonality Theorem can be written as a "dot" product:

$$\boldsymbol{A}^k \cdot \boldsymbol{A}^{k'*} = |G|\delta_{k,k'} \,.$$

This states that the irreducible representations of an Abelian group are *orthogonal* vectors in this $|G|$-dimensional space. Since there can be at most $|G|$ such vectors, the number of irreducible representations of an Abelian group is less than or equal to the order of the group.

4. (a) From Problem 2 of Problem Set 5, the irreducible representations of the three element group are:

	e	a	b
Γ_1	1	1	1
Γ_2	1	$e^{2\pi i/3}$	$e^{4\pi i/3}$
Γ_2	1	$e^{4\pi i/3}$	$e^{2\pi i/3}$

In the notation of Problem 3, we have

$$\boldsymbol{A}^1 = (1,1,1), \quad \boldsymbol{A}^2 = (1, e^{2\pi i/3}, e^{4\pi i/3}), \quad \boldsymbol{A}^3 = (1, e^{4\pi i/3}, e^{2\pi i/3}) \,.$$

Note that some of these entries are complex. Thus, the distinct inner products between these vectors are

$$\boldsymbol{A}^1 \cdot \boldsymbol{A}^{1*} = 1 + 1 + 1 = 3 \,,$$

$$
\begin{aligned}
\boldsymbol{A}^2 \cdot \boldsymbol{A}^{2*} &= 1+1+1=3\,,\\
\boldsymbol{A}^3 \cdot \boldsymbol{A}^{3*} &= 1+1+1=3\,,\\
\boldsymbol{A}^1 \cdot \boldsymbol{A}^{2*} &= 1+\mathrm{e}^{-2\pi i/3}+\mathrm{e}^{-4\pi i/3}=0\,,\\
\boldsymbol{A}^1 \cdot \boldsymbol{A}^{3*} &= 1+\mathrm{e}^{-4\pi i/3}+\mathrm{e}^{-2\pi i/3}=0\,,\\
\boldsymbol{A}^2 \cdot \boldsymbol{A}^{3*} &= 1+\mathrm{e}^{-2\pi i/3}+\mathrm{e}^{2\pi i/3}=0\,,
\end{aligned}
$$

all of which are consistent with the Great Orthogonality Theorem.

(b) In view of the fact that there are 3 mutually orthogonal vectors, there can be no additional irreducible representations of this group.

(c) For cyclic groups of order $|G|$, we determined that the irreducible representations were based on the $|G|$th roots of unity (Problem 3, Problem Set 5). Since this produces $|G|$ distinct irreducible representations, our procedure yields *all* of the irreducible representations of any cyclic group.

5. Every irreducible representation of an Abelian group is one-dimensional. Moreover, since every one of these representations is either a homomorphism or isomorphism of the group, with the operation in the representation being ordinary multiplication, the identity always corresponds to unity (Problem 9, Problem Set 3). Now, the order n of a group element g is the smallest integer for which

$$g^n = e\,.$$

For every element in any group $1 \le n \le |G|$. This relationship *must* be preserved by the irreducible representation. Thus, if A_g^k is the entry corresponding to the element g in the kth irreducible representation, then

$$(A_g^k)^n = 1\,,$$

i.e., A_g^k is the nth root of unity:

$$A_g^k = \mathrm{e}^{2m\pi i/n}, \quad m = 0, 1, \ldots, n-1\,.$$

The modulus of each of these quantities is clearly unity, so the modulus of *every* entry in the irreducible representations of an Abelian group is unity.

This is consistent with the Great Orthogonality Theorem when applied to a given representation (cf. Problem 3):

$$\sum_\alpha A_\alpha^k A_\alpha^{k^*} = \sum_\alpha \left|A_\alpha^k\right|^2 = |G|\,. \tag{1}$$

Group Theory

Problem Set 7 November 20, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1.* Identify the 12 symmetry operations of a regular hexagon.

2. Show that elements in the same class of a group must have the same order.

3.* Identify the 6 classes of this group.

Hint: You do not need to compute the conjugacy classes explicitly. Refer to the discussion for the group S_3 in Example 2.9 and use the fact that elements in the same class have the same order.

4.* How many irreducible representations are there and what are their dimensions?

5.* Construct the character table of this group by following the procedure outlined below:

(a) Enter the characters for the identical and "parity" representations. As in the case of S_3, the characters for the parity representation are either $+1$ or -1, depending on whether or not the the operation preserves the "handedness" of the coordinate system.

(b) Enter the characters for the "coordinate" representation obtained from the action on (x, y) for each group operation. Note that the character is the same for elements in the same class.

(c) Use the products $C_3C_3^2 = E$ and $C_3^3 = E$ to identify the characters for all one-dimensional irreducible representations for the appropriate classes. The meaning of the notation C_n^m for rotations is discussed in Section 5.4.

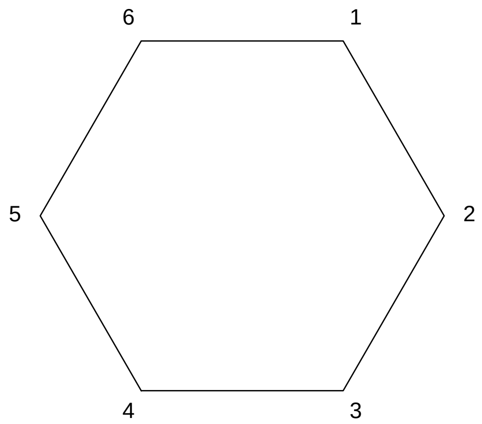
(d) Use the result of (c) and the products $C_6C_3 = C_2$ to deduce that the characters for the class of C_6 and those for the class of C_2 are the same. Then, use the orthogonality of the *columns* of the character table to compute these characters.

(e) Use the appropriate orthogonality relations for characters to compute the remaining entries of the character table.

Group Theory

Solutions to Problem Set 7 November 30, 2001

1. A regular hexagon is shown below:

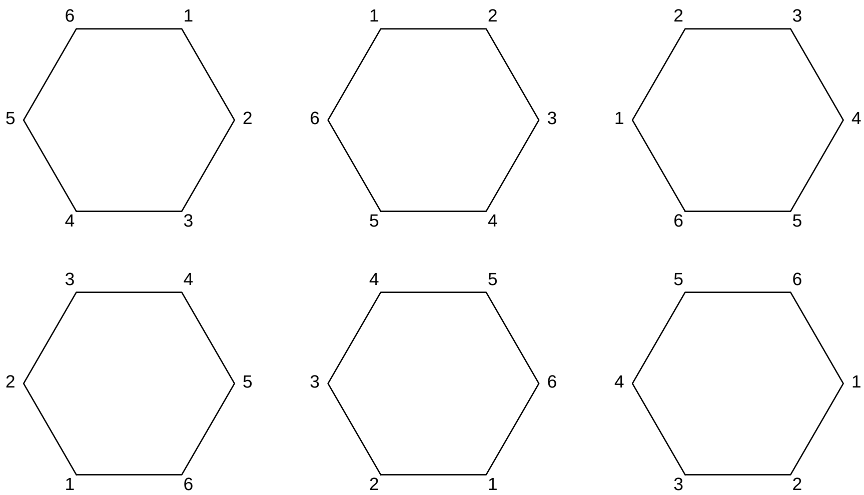


The following notation will be used for the symmetry operations of this hexagon:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \end{pmatrix},$$

where the first row corresponds to the reference order of the vertices shown in the diagram and the a_i denote the number at the ith vertex *after* the transformation of the hexagon.

The symmetry operations of this hexagon consist of the identity, rotations by angles of $\frac{1}{3}n\pi$ radians, where $n = 1, 2, 3, 4, 5$, three mirror planes which pass through opposite *faces* of the hexagon, and three mirror planes which pass through opposite *vertices* of the hexagon. For the identity and the rotations, the effect on the hexagon is



These operations correspond to

$$E = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix},$$

$$C_6 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 1 \end{pmatrix},$$

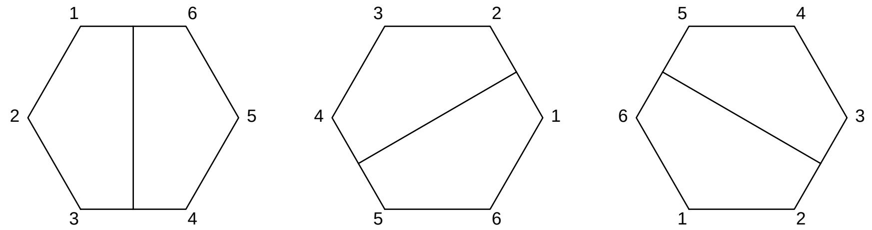
$$C_6^2 = C_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 1 & 2 \end{pmatrix},$$

$$C_6^3 = C_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 1 & 2 & 3 \end{pmatrix},$$

$$C_6^4 = C_3^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 1 & 2 & 3 & 4 \end{pmatrix},$$

$$C_6^5 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

The three mirror planes which pass through opposite *faces* of the hexagon are



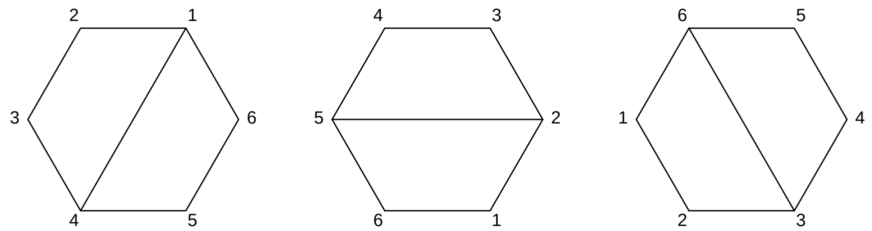
which correspond to

$$\sigma_{v,1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix} ,$$

$$\sigma_{v,2} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 6 & 5 & 4 & 3 \end{pmatrix} ,$$

$$\sigma_{v,3} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 2 & 1 & 6 & 5 \end{pmatrix} .$$

Finally, the three mirror planes which pass through opposite *vertices* of the hexagon are



These operations correspond to

$$\sigma_{d,1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 5 & 4 & 3 & 2 \end{pmatrix} ,$$

$$\sigma_{d,2} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 6 & 5 & 4 \end{pmatrix} ,$$

$$\sigma_{d,3} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 4 & 3 & 2 & 1 & 6 \end{pmatrix} .$$

That these 12 elements do, in fact, form a group is straightforward to verify. The standard notation for this group is C_{6v}.

2. The order n of an element a of a group is defined as the smallest integer such that $a^n = e$. If group elements a and b are in the same class, then there is an element g in the group such that

$$b = g^{-1} a g .$$

The m-fold product of b is then given by

$$b^m = \underbrace{(g^{-1}ag)(g^{-1}ag)\cdots(g^{-1}ag)}_{m\,\text{factors}} = g^{-1}a^m g .$$

If this is equal to the unit element e, we must have

$$g^{-1}a^m g = e ,$$

or,

$$a^m = e .$$

The smallest value of m for which this equality can be satisfied is, by definition, n, the order of a. Hence, two elements in the same class have the same order.

3. For two elements a and b of a group to be in the same class, there must be another group element such that $b = g^{-1}ag$. If the group elements are coordinate transformations, then elements in the same class correspond to the same type of operation, but in coordinate systems related by symmetry operations. This fact, together with the result of Problem 2, allows us to determine the classes of the group of the hexagon.

The identity, as always, is in a class by itself. Although all of the rotations are the same type of operation, not all of these rotations have the same orders: C_6 and C_6^5 have order 6, C_3 and C_3^2 have order 3, and C_2 has order 2. Thus, the 5 rotations belong to *three* different classes.

The two types of mirror planes, $\sigma_{v,i}$ and $\sigma_{d,i}$, *must* belong to different classes since there is no group operation which will transform any of the $\sigma_{v,i}$ to any of the $\sigma_{d,i}$. To do so would require a rotation by an odd multiple of $\frac{1}{6}\pi$, which is not a group element. All of the $\sigma_{v,i}$ are in the same class and all of the $\sigma_{d,i}$ are in the same class, since each is the same type of operation, but in coordinate systems related by symmetry operations (one of the rotations) and, of course, they all have order 2, since each reflection plane is its own inverse.

Hence, there are six classes in this group are

$$\begin{aligned}
E &\equiv \{E\}\,,\\
2C_6 &\equiv \{C_6, C_6^5\}\,,\\
2C_3 &\equiv \{C_3, C_3^2\}\,,\\
C_2 &\equiv \{C_2\}\,,\\
3\sigma_v &\equiv \{\sigma_{v,1}, \sigma_{v,2}, \sigma_{v,3}\}\,,\\
3\sigma_d &\equiv \{\sigma_{d,1}, \sigma_{d,2}, \sigma_{d,3}\}\,.
\end{aligned}$$

4. As there are 6 classes, there are 6 irreducible representations, the dimensions of which must satisfy the sum rule

$$\sum_{k=1}^{6} d_k^2 = 12\,,$$

since $|C_{6v}| = 12$. The only positive integer solutions of this equation are

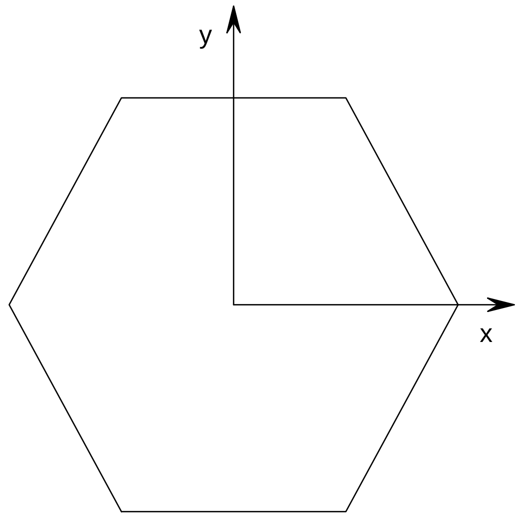
$$d_1 = 1, \quad d_2 = 1, \quad d_3 = 1, \quad d_4 = 1, \quad d_5 = 2, \quad d_6 = 2\,,$$

i.e., there are 4 one-dimensional irreducible representations and 2 two-dimensional irreducible representations.

5. (a) For the identical representation, all of the characters are 1. For the parity representation, the character is 1 for operations which preserve the parity of the coordinate system ("proper" rotations) and -1 for operations which change the parity of the coordinate system ("improper" rotations). Additionally, we can enter immediately the *column* of characters for the class of the unit element. These are equal to the dimensionality of each irreducible representation, since the unit element is the identity matrix with that dimensionality. Thus, we have the following entries for the character table of C_{6v}:

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
Γ_1	1	1	1	1	1	1
Γ_1'	1	1	1	1	-1	-1
Γ_1''	1					
Γ_1'''	1					
Γ_2	2					
Γ_2'	2					

(b) The characters for one of the two-dimensional representations of C_{6v} can be obtained by constructing matrices for operations in analogy with the procedure discussed in Section 3.2 for the equilateral triangle. One important difference here is that we require such a construction only for one element in each class (since the all matrices in a given class have the same trace). We will determine the representations of operations in each class in an (x, y) coordinate system shown below:



Thus, a rotation by an angle ϕ, denoted by $R(\phi)$, is given by the two-dimensional rotational matrix:

$$R(\phi) = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}.$$

The corresponding character $\chi(\phi)$ is, therefore, simply the sum of the diagonal elements of this matrix:

$$\chi(\phi) = 2\cos\phi\,.$$

We can now calculate the characters for each of the classes composed of rotations:

$$\begin{aligned} \chi(2C_6) &= \chi(\tfrac{1}{3}\pi) = 1\,, \\ \chi(2C_3) &= \chi(\tfrac{2}{3}\pi) = -1\,, \\ \chi(C_2) &= \chi(\pi) = -2\,. \end{aligned}$$

For the two classes of mirror planes, we need only determine the character of one element in each class, which may be chosen at our convenience. Thus, for example, since the representation of

$\sigma_{v,1}$ can be determined directly by inspection:

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

we can obtain the character of the corresponding class as

$$\chi(3\sigma_v) = 0\,.$$

Similarly, the representation of $\sigma_{d,2}$ can also be determined directly by inspection:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which yields the character

$$\chi(3\sigma_d) = 0\,.$$

We can now add the entries for this two-dimensional irreducible representations to the character table of C_{6v}:

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
Γ_1	1	1	1	1	1	1
Γ_1'	1	1	1	1	-1	-1
Γ_1''	1					
Γ_1'''	1					
Γ_2	2	1	-1	-2	0	0
Γ_2'	2					

(c) The one-dimensional irreducible representations must obey the multiplication table, since they themselves are representations of the group. In particular, given the products

$$C_3C_3^2 = E, \qquad C_3^3 = E\,,$$

if we denote by α the character of the class $2C_3 = \{C_3, C_3^2\}$, then these products require that

$$\alpha^2 = 1, \qquad \alpha^3 = 1\,,$$

respectively. Thus, we deduce that $\alpha = 1$ for *all* of the one-dimensional irreducible representations. With these additions to the character table, we have

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
Γ_1	1	1	1	1	1	1
Γ_1'	1	1	1	1	-1	-1
Γ_1''	1		1			
Γ_1'''	1		1			
Γ_2	2	1	-1	-2	0	0
Γ_2'	2					

(d) Since the character for all one-dimensional irreducible representations for the class $2C_3 = \{C_3, C_3^2\}$ is unity, the product $C_6C_3 = C_2$ requires that the characters for the classes of C_6 and C_2 are the same in these representations. Since $C_2^2 = E$, this character must be 1 or -1. Suppose we choose $\chi(2C_6) = \chi(C_2) = 1$. Then, the orthogonality of the *columns* of the character table requires that the character for the classes E and $2C_6$ are orthogonal. If we denote by β the character for the class $2C_6$ of the representation Γ_2', we require

$$(1 \times 1) + (1 \times 1) + (1 \times 1) + (1 \times 1) + (2 \times 1) + (2 \times \beta) = 0\,,$$

i.e., $\beta = -3$. But this value violates the requirement that

$$\sum_\alpha n_\alpha |\chi_\alpha|^2 = |G|\,. \tag{1}$$

Thus, we must choose $\chi(2C_6) = \chi(C_2) = -1$, and our character table becomes

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
Γ_1	1	1	1	1	1	1
Γ_1'	1	1	1	1	-1	-1
Γ_1''	1	-1	1	-1		
Γ_1'''	1	-1	1	-1		
Γ_2	2	1	-1	-2	0	0
Γ_2'	2					

(e) The characters for the classes $2C_6$, $2C_3$, and C_2 of the Γ_2' representation can now be determined by requiring the columns corresponding to these classes to be orthogonal to the column corresponding to the class of the identity. When this is done, we find that the values obtained saturate the sum rule in (1), so the characters corresponding to both classes of mirror planes in this representation must vanish. This enables to complete the entries for the Γ_2' representation:

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
Γ_1	1	1	1	1	1	1
Γ_1'	1	1	1	1	-1	-1
Γ_1''	1	-1	1	-1		
Γ_1'''	1	-1	1	-1		
Γ_2	2	1	-1	-2	0	0
Γ_2'	2	-1	1	2	0	0

The remaining entries are straightforward to calculate. The fact that each mirror reflection has order 2 means that these entries must be either $+1$ or -1. The requirement of orthogonality of columns leaves only one choice:

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
Γ_1	1	1	1	1	1	1
Γ_1'	1	1	1	1	-1	-1
Γ_1''	1	-1	1	-1	1	-1
Γ_1'''	1	-1	1	-1	-1	1
Γ_2	2	1	-1	-2	0	0
Γ_2'	2	-1	1	2	0	0

which completes the character table for C_{6v}.

Group Theory

Problem Set 8 November 27, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1. Show that if two matrices A and B are orthogonal, then their direct product $A \otimes B$ is also an orthogonal matrix.

2. Show that the trace of the direct product of two matrices A and B is the product of the traces of A and B:

$$\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\,\mathrm{tr}(B)$$

3.* Show that the direct product of groups G_a and G_b with elements $G_a = \{e, a_2, \ldots, a_{|G_a|}\}$ and $G_b = \{e, b_2, \ldots, b_{|G_b|}\}$, such that $a_i b_j = b_j a_i$ for all i and j, is a group. What is the order of this group?

4.* Use the Great Orthogonality Theorem to show that the direct product of irreducible representations of two groups is an irreducible representation of the direct product of those groups.

5.* For an n-fold degenerate set of eigenfunctions φ_i, $i, 1, 2, \ldots, n$, we showed show that the matrices $\Gamma(R_\alpha)$ generated by the group of the Hamiltonian,

$$R_\alpha \varphi_i = \sum_{j=1}^{n} \varphi_j \Gamma_{ji}(R_\alpha)$$

form a representation of that group. Show that if the φ_j are chosen to be an orthonormal set of functions, then this representation is *unitary*.

6.* The set of distinct functions obtained from a given function φ_i by operations in the group of the Hamiltonian, $\varphi_j = R_\alpha \varphi_i$, are called **partners**. Use the Great Orthogonality Theorem to show that two functions which belong to different irreducible representations or are different partners in the same unitary representation are orthogonal.

7. Consider a particle of mass m confined to a square in two dimensions whose vertices are located at $(1, 1)$, $(1, -1)$, $(-1, -1)$, and $(-1, 1)$. The potential is taken to be zero within the square and infinite at the edges of the square. The eigenfunctions φ are of the form

$$\varphi_{p,q}(x, y) \propto \left\{ \begin{array}{c} \cos(k_p x) \\ \sin(k_p x) \end{array} \right\} \left\{ \begin{array}{c} \cos(k_q y) \\ \sin(k_q y) \end{array} \right\}$$

where $k_p = \frac{1}{2}p\pi$, $k_q = \frac{1}{2}q\pi$, and p and q are positive integers. The notation above means that $\cos(k_p x)$ is taken if p is odd, $\sin(k_p x)$ is taken if p is even, and similarly for the other factor. The corresponding eigenvalues are

$$E_{p,q} = \frac{\hbar^2\pi^2}{8m}(p^2 + q^2)$$

(a) Determine the eight planar symmetry operations of a square. These operations form the group of the Hamiltonian for this problem. Assemble the symmetry operations into equivalence classes.

(b) Determine the number of irreducible representations and their dimensions for this group. Do these dimensions appear to be broadly consistent with the degeneracies of the energy eigenvalues?

(c) Determine the action of each group operation on (x, y).

Hint: This can be done by inspection.

(d) Determine the characters corresponding to the identical, parity, and coordinate representations. Using appropriate orthogonality relations, complete the character table for this group.

(e) For which irreducible representations do the eigenfunctions $\varphi_{1,1}(x, y)$ and $\varphi_{2,2}(x, y)$ form bases?

(f) For which irreducible transformation do the eigenfunctions $\varphi_{1,2}(x, y)$ and $\varphi_{2,1}(x, y)$ form a basis?

(g) What is the degeneracy corresponding to $(p = 6, q = 7)$ and $(p = 2, q = 9)$? Is this a normal or accidental degeneracy?

(h) Are there eigenfunctions which form a basis for each of the irreducible representations of this group?

8.* Consider the regular hexagon in Problem Set 7. Suppose there is a vector perturbation, i.e., a perturbation that transforms as (x, y, z). Determine the selection rule associated with an initial state that transforms according to the "parity" representation.

Hint: The reasoning for determining the irreducible representations associated with (x, y, z) is analogous to that used in Section 6.6.2 for the equilateral triangle.

Group Theory

Solutions to Problem Set 8 December 7, 2001

1. A matrix A is said to be orthogonal if its matrix elements a_{ij} satisfy the following relations:

$$\sum_i a_{ij}a_{ik} = \delta_{j,k}, \qquad \sum_j a_{ij}a_{kj} = \delta_{ij}\,, \tag{1}$$

i.e., the rows and columns are orthogonal vectors. This ensures that $A^{\rm t}A = AA^{\rm t} = I$.

The direct product C of two matrices A and B, denoted by $C = A \otimes B$, is given in terms of matrix elements by

$$c_{ik;jl} = a_{ij}a_{kl}\,.$$

If A and B are orthogonal matrices, then we can show that C is also an orthogonal matrix by verifying the relations in Eq. (1). The first of these relations is

$$\sum_{ik} c_{ik;jl}c_{ik;j'l'}$$

$$= \sum_{ik} a_{ij}a_{kl}a_{ij'}a_{kl'} = \Big(\sum_i a_{ij}a_{ij'}\Big)\Big(\sum_k a_{kl}a_{kl'}\Big) = \delta_{j,j'}\delta_{l,l'}\,,$$

where the last step follows from the first of Eqs. (1). The second orthogonality relation is

$$\sum_{jl} c_{ik;jl}c_{i'k';jl}$$

$$= \sum_{jl} a_{ij}a_{kl}a_{i'j}a_{k'l} = \Big(\sum_j a_{ij}a_{i'j}\Big)\Big(\sum_l a_{kl}a_{k'l}\Big) = \delta_{i,i'}\delta_{k,k'}$$

where the last step follows from the second of Eqs. (1). Thus, we have shown that the direct product of two orthogonal matrices is also an orthogonal matrix.

2. The direct product of two matrices A and B with matrix elements a_{ij} and b_{ij} is

$$c_{ik,jl} = a_{ij} b_{kl} \,.$$

The trace of the direct product $A \otimes B$ is obtained by setting $j = i$ and $l = k$ and summing over i and k:

$$\mathrm{tr}(A \otimes B) = \sum_{ik} c_{ik,ik} = \sum_{ik} a_{ii} b_{kk} = \sum_{i} a_{ii} \sum_{k} b_{kk} = \mathrm{tr}(A)\,\mathrm{tr}(B)\,,$$

which is the product of the traces of A and B.

3. We have two groups G_a and G_b with elements

$$G_a = \{e_a, a_2, a_3, \ldots, a_{|G_a|}\}$$

and

$$G_b = \{e_b, b_2, b_3, \ldots, b_{|G_b|}\}\,,$$

such that $a_i b_j = b_j a_i$ for all i and j. We are using a notation where it is understood that $a_1 = e_a$ and $b_1 = e_b$. The direct product $G_a \otimes G_b$ of these groups is the set obtained by forming the product of every element of G_a with every element of G_b:

$$G_a \otimes G_b = \{e, a_2, a_3, \ldots, a_{n_a}, b_2, b_3, \ldots, b_{n_b}, \ldots, a_i b_j, \ldots\}\,.$$

To show that $G_a \otimes G_b$ is a group, we must demonstrate that these elements fulfill each of the four requirements in Sec. 2.1.

Closure. The product of two elements $a_i b_j$ and $a_{i'} b_{j'}$ is given by

$$(a_i b_j)(a_{i'} b_{j'}) = (a_i a_{i'})(b_j b_{j'}) = a_k b_l\,,$$

where the first step follows from the commutativity of elements between the two groups and the second step from the group property of G_a and G_b.

Associativity. The associativity of the composition law follows from

$$\begin{aligned}(a_i b_{i'} a_j b_{j'}) a_k b_{k'} &= \Big[(a_i a_j) a_k\Big]\Big[(b_{i'} b_{j'}) b_{k'}\Big] \\ &= \Big[a_i (a_j a_k)\Big]\Big[b_{i'} (b_{j'} b_{k'})\Big] \\ &= a_i b_{i'} (a_j b_{j'} a_k b_{k'})\,,\end{aligned}$$

since associativity holds for G_a and G_b separately.

Unit Element. The unit element e for the direct product group is $e_a e_b = e_b e_a$, since

$$(a_i b_j)(e_a e_b) = (a_i e_a)(b_j e_b) = (e_a a_i)(e_b b_j) = (e_a e_b)(a_i b_j)\,.$$

Inverse. Finally, the inverse of each element $a_i b_j$ is $a_i^{-1} b_j^{-1}$ because

$$(a_i b_j)(a_i^{-1} b_j^{-1}) = (a_i a_i^{-1})(b_j b_j^{-1}) = e_a e_b$$

and

$$(a_i^{-1} b_j^{-1})(a_i b_j) = (a_i^{-1} a_i)(b_j^{-1} b_j) = e_a e_b\,.$$

Thus, we have shown that the direct product of two groups is itself a group. Since the elements of this group are obtained by taking all products of elements from G_a and G_b, the order of this group is $|G_a||G_b|$.

4. Suppose we have an irreducible representation for each of two groups G_a and G_b. We denote these representations, which may be of different dimensions, by $A(a_i)$ and $A(b_j)$, and their matrix elements by $A(a_i)_{ij}$ and $A(b_j)_{ij}$. Since these representations are irreducible, they satisfy the Great Orthogonality Theorem:

$$\sum_{a_i} A(a_i)^*_{ij} A(a_i)_{i'j'} = \frac{|G_a|}{d_a} \delta_{i,i'} \delta_{j,j'}\,,$$

$$\sum_{b_j} A(b_j)^*_{ij} A(b_j)_{i'j'} = \frac{|G_b|}{d_b} \delta_{i,i'} \delta_{j,j'}\,,$$

where d_a and d_b are the dimensions of the irreducible representations of G_a and G_b, respectively. A representation of the direct product of two groups, denoted by $A(a_i b_j)$, is obtained from the direct product of representations of each group:

$$A(a_i b_j)_{ik;jl} = A(a_i)_{ij} A(b_j)_{kl} \,.$$

The sum in the Great Orthogonality Theorem for the direct product representation is

$$\begin{aligned}
&\sum_{a_i}\sum_{b_j} A(a_i b_j)^*_{ik;jl} A(a_i b_j)_{i'k';j'l'} \\
&= \sum_{a_i}\sum_{b_j} A(a_i)^*_{ij} A(b_j)^*_{kl} A(a_i)_{i'j'} A(b_j)_{k'l'} \\
&= \underbrace{\Big[\sum_{a_i} A(a_i)^*_{ij} A(a_i)_{i'j'}\Big]}_{\frac{|G_a|}{d_a}\delta_{i,i'}\delta_{j,j'}} \underbrace{\Big[\sum_{b_j} A(b_j)^*_{kl} A(b_j)_{k'l'}\Big]}_{\frac{|G_b|}{d_b}\delta_{k,k'}\delta_{l,l'}} \\
&= \left(\frac{|G_a||G_b|}{d_a d_b}\right)\delta_{i,i'}\delta_{k,k'}\delta_{j,j'}\delta_{l,l'} \,.
\end{aligned}$$

This shows that this direct product representation is, in fact, irreducible. It has dimensionality $d_a d_b$ and the order of the direct product is, of course, $|G_a| \times |G_b|$.

5. If the φ_i are orthonormal, and if this property is required to be preserved by the group of the Hamiltonian (as it must, to conserve probability), then, in Dirac notation, we have

$$\begin{aligned}
(i,j) &\equiv \int \varphi_i(x)^* \varphi_j(x)\,\mathrm{d}x = \int [R\varphi_i(x)]^* R\varphi_j(x)\,\mathrm{d}x \\
&= (i|R^\dagger R|j) = \delta_{i,j} \,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
(i|R^{\dagger}R|j) &= \sum_{k,l}(k,l)\Gamma(R)^{*}_{ki}\Gamma(R)_{lj} \\
&= \sum_{k}\Gamma(R)^{*}_{ki}\Gamma(R)_{kj} \\
&= \sum_{k}\left[\Gamma(R)^{\dagger}\right]_{ik}\Gamma(R)_{kj} \\
&= \left[\Gamma(R)^{\dagger}\Gamma(R)\right]_{ij},
\end{aligned}$$

i.e., when written in matrix notation,

$$\Gamma(R)^{\dagger}\Gamma(R) = I\,.$$

Thus, the matrix representation is unitary.

6. We again use Dirac notation to signify basis functions φ_i and φ_j corresponding to irreducible representations n and n', respectively: $|n,i)$ and $|n',j)$. Then, the operations R in the group of the Hamiltonian applied to these functions yield

$$\begin{aligned}
R|n,i) &= \sum_{k}\Gamma^{(n)}(R)_{ki}|n,k)\,, \\
R|n',j) &= \sum_{l}\Gamma^{(n')}(R)_{lj}|n',l)\,.
\end{aligned}$$

Since the operators and their representations are unitary,

$$(n',j|R^{\dagger} = (n',j|R^{-1} = \sum_{l}\Gamma^{(n')}(R)^{*}_{lj}(n',l|\,,$$

we have

$$\begin{aligned}
(n',j|R^{-1}R|n,i) &= (n',j|n,i) \\
&= \sum_{kl}\Gamma^{(n')}(R)^{*}_{kj}\Gamma^{(n)}(R)_{li}(n',k|n,l)\,.
\end{aligned}$$

If we now sum both sides of this equation over the elements of the group of the Hamiltonian, and invoke the Great Orthogonality Theorem, we obtain

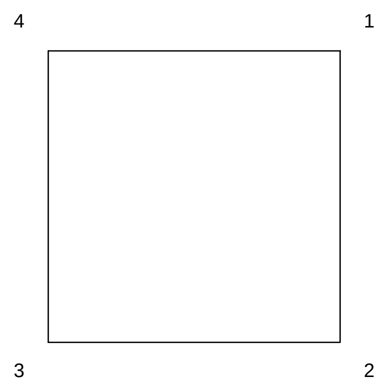
$$\begin{aligned}\sum_R (n', j|n, i) &= |G|(n', j|n, i) \\ &= \sum_{kl} \underbrace{\Big[\sum_R \Gamma^{(n')}(R)^*_{kj}\Gamma^{(n)}(R)_{li} \Big]}_{\frac{|G|}{d_n}\delta_{n,n'}\delta_{k,l}\delta_{i,j}} (n', k|n, l) \\ &= |G|\delta_{n,n'}\delta_{i,j}(n', k|n, k)\,,\end{aligned}$$

where $|G|$ is the order of the group of the Hamiltonian and d_n is the dimension of the nth irreducible representation. Therefore,

$$(n', j|n, i) = \delta_{n,n'}\delta_{i,j}\,,$$

since $(n, k|n, k) = 1$.

7. (a) A square is shown below:



In analogy with the procedure described in Problem Set 7, we will use the following notation for the symmetry operations of

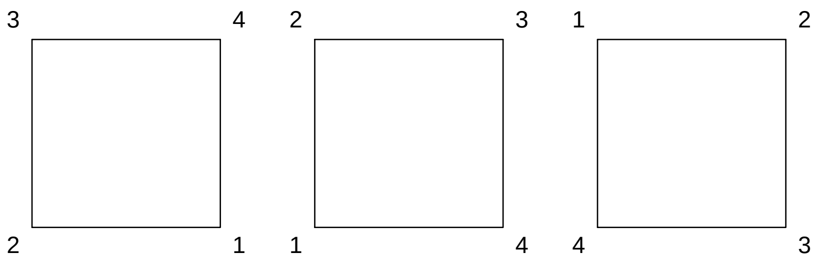
this hexagon:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ a_1 & a_2 & a_3 & a_4 \end{pmatrix},$$

where the a_i denote the number at the ith vertex after the transformation of the hexagon given in the indicated reference order. Thus, the identity operation, which identifies the reference order of the vertices, corresponds to

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

The symmetry operations on this square consist of the identity, rotations by angles of $\frac{1}{2}n\pi$ radians, for $n = 1, 2, 3$, two mirror planes which pass through opposite *faces* of the square, and two mirror planes which pass through opposite *vertices* of the square. For the rotations, the effect on the square is



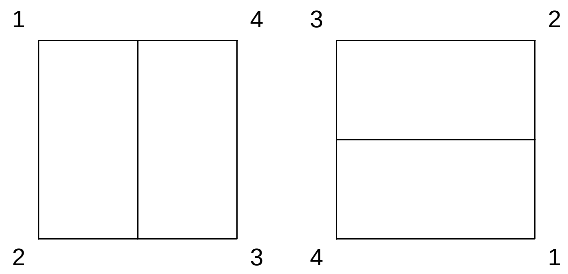
These operations correspond to

$$C_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix},$$

$$C_4^2 = C_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix},$$

$$C_4^3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}.$$

The two mirror planes which pass through opposite *faces* of the square are

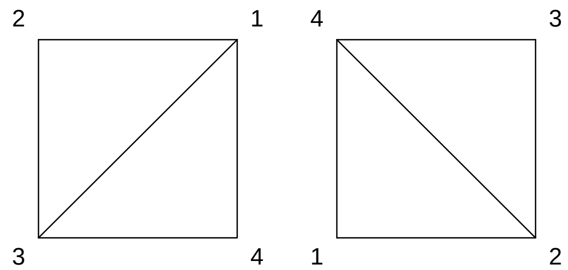


which correspond to

$$\sigma_{v,1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} ,$$

$$\sigma_{v,2} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} .$$

Finally, the two mirror planes which pass through opposite *vertices* of the square are



These operations correspond to

$$\sigma'_{v,1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} ,$$

$$\sigma'_{v,2} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} .$$

Elements in the same equivalence class must have the same order and correspond to the same "type" of operation. Thus, there are *five* equivalence classes of this group:

$$\{E\}, \quad \{C_2 = C_4^2\}, \quad \{2C_4\}, \quad \{2\sigma_v\}, \quad \{2\sigma_v'\}\,.$$

Note that, as in the case of the regular hexagon (Problem Set 7), all of the rotations need not belong to the same class, despite being the same "type" of operation because they must also have the same order.

(b) The order of this group is 8 and there are 5 equivalence classes. Thus, there must be five irreducible representations whose dimensionalities must satisfy

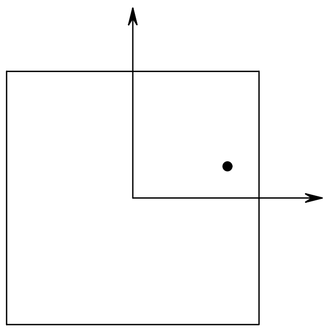
$$d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2 = 8\,.$$

The *only* solution of this equation (with positive integer values for the d_k) is

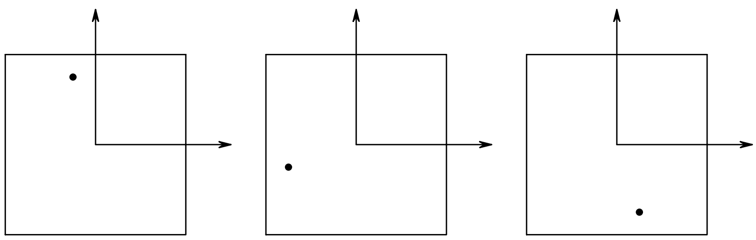
$$d_1 = 1, \quad d_2 = 1, \quad d_3 = 1, \quad d_4 = 1, \quad d_5 = 2\,.$$

These dimensionalities imply that the energy levels for a Hamiltonian with this symmetry are either nondegenerate or are two-fold degenerate. From the expression given for the energy eigenvalues, we see immediately that the energy eigenvalues with $p = q$ are non-degenerate, and those with $p \neq q$ are two-fold degenerate (but see below). Thus, the dimensions of the irreducible representations are consistent with these degeneracies.

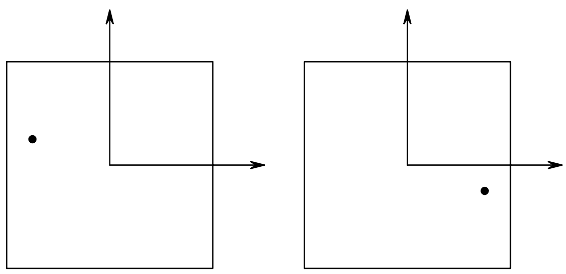
(c) The simplest way to obtain a two-dimensional representation of this group is to consider the action of each group element on some generic point (x, y). Then the action on this point of each of the operations given above can be determined by inspection. We begin with the figure below:



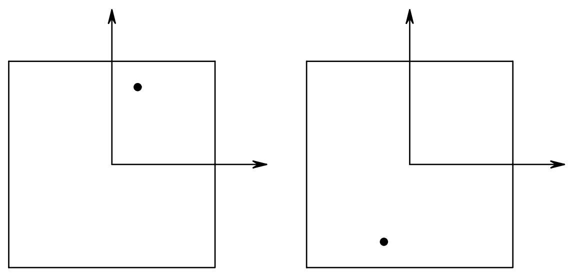
From the diagrammatic representation of each symmetry operation, we will be able to determine the corresponding representation, simply by inspection. The action on this point by the three rotations acn be represented as



These rotations are thus seen to transform the point (x, y) into $(-y, x)$, $(-x, -y)$, and $(y, -x)$, respectively. The two reflections that pass through the center of faces are



so they transform the (x, y) into $(-x, y)$ and $(x, -y)$, respectively. Finally, for the two reflection planes which pass through vertices,



which transform the point (x, y) into (y, x) and $(-y, -x)$, respectively. These transformations enable us to construct the characters corresponding to the "coordinate" representation. Then, together with the identical and parity representations, we have the following entries of the character table for this group:

	E	C_2	$2C_4$	$2\sigma_v$	$2\sigma'_v$
Γ_1	1	1	1	1	1
Γ'_1	1	1	1	-1	-1
Γ''_1	1				
Γ'''_1	1				
Γ_2	2	-2	0	0	0

The group multiplication table and the orthogonality of *columns* allows us to immediately complete the entries for the classes $\{C_2\}$ and $\{2C_4\}$:

	E	C_2	$2C_4$	$2\sigma_v$	$2\sigma'_v$
Γ_1	1	1	1	1	1
Γ'_1	1	1	1	-1	-1
Γ''_1	1	1	-1		
Γ'''_1	1	1	-1		
Γ_2	2	-2	0	0	0

The remaining four entries can be determined from the orthogonality of either rows or columns and again invoking the group multiplication table:

	E	C_2	$2C_4$	$2\sigma_v$	$2\sigma_v'$
Γ_1	1	1	1	1	1
Γ_1'	1	1	1	-1	-1
Γ_1''	1	1	-1	1	-1
Γ_1'''	1	1	-1	-1	1
Γ_2	2	-2	0	0	0

(e) The eigenfunctions $\varphi_{1,1}(x,y)$ and $\varphi_{2,2}(x,y)$ are given by

$$\varphi_{1,1}(x,y) \propto \cos(\tfrac{1}{2}\pi x)\cos(\tfrac{1}{2}\pi y)$$

and

$$\varphi_{2,2}(x,y) \propto \sin(\pi x)\sin(\pi y)\,.$$

Since $\varphi_{1,1}(x,y)$ is invariant under the interchange of x and y and under changes in their signs, it transforms according to the **identical** representation. However, although $\varphi_{2,2}(x,y)$ is invariant under the interchange of x and y, each sine factor changes sign if their argument changes sign. Thus, this eigenfunction transforms according to the **parity** representation.

(f) The (degenerate) eigenfunctions $\varphi_{1,2}(x,y)$ and $\varphi_{2,1}(x,y)$ are

$$\begin{pmatrix}\varphi_{1,2}\\ \varphi_{2,1}\end{pmatrix} \propto \begin{bmatrix}\cos(\tfrac{1}{2}\pi x)\sin(\pi y)\\ \sin(\pi x)\cos(\tfrac{1}{2}\pi y)\end{bmatrix}\,.$$

The transformation properties of these eigenfunctions can be determined from the results of part (c). This yields the following matrix representation of each symmetry operation:

$$E=\begin{pmatrix}1&0\\0&1\end{pmatrix}, \qquad C_4=\begin{pmatrix}0&1\\-1&0\end{pmatrix}, \qquad C_4^2=\begin{pmatrix}-1&0\\0&-1\end{pmatrix},$$

$$C_4^3=\begin{pmatrix}0&-1\\1&0\end{pmatrix}, \qquad \sigma_{v,1}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}, \qquad \sigma_{v,2}=\begin{pmatrix}-1&0\\0&1\end{pmatrix},$$

$$\sigma_{v,1}'=\begin{pmatrix}0&1\\1&0\end{pmatrix}, \qquad \sigma_{v,2}'=\begin{pmatrix}0&-1\\-1&0\end{pmatrix}.$$

This produces the following characters:

$$\{E\} = 2, \quad \{C_2\} = -2, \quad \{2C_4\} = 0, \quad \{2\sigma_v\} = 0, \quad \{2\sigma_v'\} = 0\,,$$

which are the characters of the two-dimensional irreducible representation Γ_2, which is the “coordinate” irreducible representation.

(g) We have that the energies $E_{6,7}$ and $E_{2,9}$ are given by

$$E_{6,7} = E_{7,6} = \frac{\hbar^2\pi^2}{8m}(6^2 + 7^2) = 85\frac{\hbar^2\pi^2}{8m}$$

and

$$E_{2,9} = E_{9,2} = \frac{\hbar^2\pi^2}{8m}(2^2 + 9^2) = 85\frac{\hbar^2\pi^2}{8m}\,,$$

so this energy is *fourfold* degenerate. However, since the group operations have the effect of interchanging x and y with possible changes of sign, the eigenfunctions $\varphi_{6,7}$ and $\varphi_{7,6}$ are transformed only between one another, and the eigenfunctions $\varphi_{2,9}$ and $\varphi_{9,2}$ are transformed only between one another. In other words, this fourfold degeneracy is **accidental**, resulting only from the numerical coincidence of the energies of two twofold-degenerate states.

(h) We have already determined that $\varphi_{p,p}$ with p even transforms according to the identical representation, while if p is odd, $\varphi_{p,p}$ transforms according to the parity representation. Moreover, the pair of eigenfunctions $\varphi_{p,q}$ where p is even and q is odd transforms according to the coordinate representation.

Consider now the case where the eigenfunctions are of the form $\varphi_{p,q}$ where *both* p and q are even. The matrices corresponding to

the symmetry operations are

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad C_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad C_4^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$C_4^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_{v,1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_{v,2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\sigma'_{v,1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma'_{v,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The corresponding characters are

$$\{E\} = 2, \quad \{C_2\} = 2, \quad \{2C_4\} = 0, \quad \{2\sigma_v\} = 2, \quad \{2\sigma'_v\} = 0\,.$$

This representation must be *reducible*, since its characters do not correspond to those of any of the irreducible representations in the table determined in Part (d). A straightforward application of the Decomposition Theorem (or simple inspection) shows that this representation is the direct sum of the Γ_1 and Γ_1'' irreducible representations. This means that there is a linear combination of these eigenfunctions that diagonalizes the matrices corresponding to each symmetry operation of this group.

For the eigenfunctions of the form $\varphi_{p,q}$ where *both* p and q are odd, the matrices corresponding to the symmetry operations are

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad C_4 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \qquad C_4^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$C_4^3 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad \sigma_{v,1} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_{v,2} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\sigma'_{v,1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma'_{v,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The characters are now

$$\{E\} = 2, \quad \{C_2\} = 2, \quad \{2C_4\} = 0, \quad \{2\sigma_v\} = -2, \quad \{2\sigma'_v\} = 0\,.$$

which correspond to a *reducible* representation composed of the direct sum of the Γ_1' and Γ_1''' irreducible representations. Thus, *all* of the irreducible representations occur in the eigenfunctions of the two-dimensional square well.

8. The character table of the regular hexagon is reproduced below:

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
Γ_1	1	1	1	1	1	1
Γ_1'	1	1	1	1	-1	-1
Γ_1''	1	-1	1	-1	1	-1
Γ_1'''	1	-1	1	-1	-1	1
Γ_2	2	1	-1	-2	0	0
Γ_2'	2	-1	1	2	0	0

A transformation properties of a vector perturbation can be deduced in a manner analogous to that for the equilateral triangle (Section 6.6.2). Applying each symmetry operation to $\boldsymbol{r} = (x, y, z)$ produces a *reducible* representation because these operations are either rotations or reflections through vertical planes. Thus, the z axis is invariant under every symmetry operation of this group which. Together with the fact that an (x, y) basis generates the two-dimensional irreducible representation Γ_2 [Problem 5(b), Problem Set 7], yields

$$\Gamma' = \Gamma_1 \oplus \Gamma_2 \,.$$

The corresponding characters are

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
$\Gamma_1 \oplus \Gamma_2$	3	2	0	-1	1	1

to determine the selection rule for an initial state that transforms according to the parity representation (Γ_1', we must calculated

$$\Gamma_1' \otimes \Gamma' = \Gamma_1' \otimes (\Gamma_1 \oplus \Gamma_2) \,.$$

The characters associated with this operation are

C_{6v}	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$
$\Gamma_1' \otimes (\Gamma_1 \oplus \Gamma_2)$	3	2	0	-1	-1	-1

Finally, either by inspection, or by applying the decomposition theorem, we find that

$$\Gamma_1' \otimes (\Gamma_1 \oplus \Gamma_2) = \Gamma_1' \oplus \Gamma_2 \,,$$

so transitions between states that transform according to the parity representation and any states other than those that transform as the parity or coordinate representations are forbidden by symmetry.

Group Theory

Problem Set 9 December 4, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1.* Consider the group $\mathrm{O}(n)$, the elements of which preserve the Euclidean length in n dimensions:

$$x_1'^2 + x_2'^2 + \cdots + x_n'^2 = x_1^2 + x_2^2 + \cdots + x_n^2 \,.$$

Show that these transformations have $\frac{1}{2}n(n-1)$ free parameters.

2. The condition that the Euclidean length is preserved in two dimensions, $x'^2 + y'^2 = x^2 + y^2$, was shown in lectures to require that

$$a_{11}^2 + a_{21}^2 = 1, \qquad a_{11}a_{12} + a_{21}a_{22} = 0, \qquad a_{12}^2 + a_{22}^2 = 1 \,.$$

Show that these requirements imply that

$$(a_{11}a_{22} - a_{12}a_{21})^2 = 1 \,.$$

3. Rotations in two dimensions can be parametrized by

$$R(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} .$$

Show that

$$R(\varphi_1 + \varphi_2) = R(\varphi_1)R(\varphi_2)$$

and, hence, deduce that this group is Abelian.

4.* We showed in lectures that a rotation $R(\varphi)$ by an angle φ in two dimensions can be written as

$$R(\varphi) = \mathrm{e}^{\varphi X} \,,$$

where

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} .$$

Verify that

$$\mathrm{e}^{\varphi X} = I\cos\varphi + X\sin\varphi \,,$$

where I is the two-dimensional unit matrix. This shows that $\mathrm{e}^{\varphi X}$ is the rotation matrix in two dimensions.

5.* Consider the two-parameter group

$$x' = ax + b$$

Determine the infinitesimal operators of this group.

6.* Consider the group $C_{\infty v}$ which contains, in addition to all two-dimensional rotations, a reflection plane, denoted by σ_v in, say, the x-z plane. Is this group Abelian? What are the equivalence classes of this group?

Hint: Denoting reflection in the x–z plane by S, show that $SR(\varphi)S^{-1} = R(-\varphi)$.

7. By extending the procedure used in lectures for SO(3), show that the infinitesimal generators of SO(4), the group of proper rotations in four dimensions which leave the quantity $x^2 + y^2 + z^2 + w^2$ invariant, are

$$A_1 = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}, \qquad A_2 = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}, \qquad A_3 = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$$

$$B_1 = x\frac{\partial}{\partial w} - w\frac{\partial}{\partial x}, \qquad B_2 = y\frac{\partial}{\partial w} - w\frac{\partial}{\partial y}, \qquad B_3 = z\frac{\partial}{\partial w} - w\frac{\partial}{\partial z}$$

8. Show that the commutators of the generators obtained in Problem 7 are

$$[A_i, A_j] = \varepsilon_{ijk}A_k, \qquad [B_i, B_j] = \varepsilon_{ijk}A_k, \qquad [A_i, B_j] = \varepsilon_{ijk}B_k$$

9. Show that by making the linear transformation of the generators in Problem 7 to

$$J_i = \tfrac{1}{2}(A_i + B_i), \qquad K_i = \tfrac{1}{2}(A_i - B_i)$$

the commutators become

$$[J_i, J_j] = \varepsilon_{ijk}J_k, \qquad [K_i, K_j] = \varepsilon_{ijk}K_k, \qquad [J_i, K_j] = 0$$

This shows that locally SO(4)=SO(3)$\otimes$SO(3).

Group Theory

Solutions to Problem Set 9 December 14, 2001

1. The Lie group GL(n,R) has n^2 parameters, because the transformations can be represented as $n \times n$ matrices (with real entries). The requirement that the Euclidean length dimensions be preserved by such a transformation leads to the requirement that,

$$x_1'^2 + x_2'^2 + \cdots + x_n'^2 = x_1^2 + x_2^2 + \cdots + x_n^2\,.$$

Proceeding as in Sec. 7.2, we note that there are n conditions from the requirement that the coefficients of x_i, $i = 1, 2, \ldots, n$ be equal to unity. Then, there are $\frac{1}{2}n(n-1)$ conditions from the requirement that the coefficients of the *unique* products $x_i x_j$, $i \neq j$ vanish. Thus, beginning with n free parameters for GL(n,R), there are

$$n^2 - n - \tfrac{1}{2}n(n-1) = n^2 - n - \tfrac{1}{2}n^2 + \tfrac{1}{2}n = \tfrac{1}{2}n(n-1)$$

free parameters for O(n).

2. Beginning with the three conditions

$$a_{11}^2 + a_{21}^2 = 1, \qquad a_{11}a_{12} + a_{21}a_{22} = 0, \qquad a_{12}^2 + a_{22}^2 = 1\,,$$

we take the product of the first by the third equations and subtract the square of the second equation to obtain

$$\begin{aligned}
&(a_{11}^2 + a_{21}^2)(a_{12}^2 + a_{22}^2) - (a_{11}a_{12} + a_{21}a_{22})^2 \\
&= a_{11}^2 a_{12}^2 + a_{11}^2 a_{22}^2 + a_{21}^2 a_{12}^2 + a_{21}^2 a_{22}^2 - a_{11}^2 a_{12}^2 \\
&\qquad -2a_{11}a_{12}a_{21}a_{22} - a_{21}^2 a_{22}^2 \\
&= a_{11}^2 a_{22}^2 + a_{21}^2 a_{12}^2 - 2a_{11}a_{12}a_{21}a_{22} \\
&= (a_{11}a_{22} - a_{12}a_{21})^2 \\
&= 1\,.
\end{aligned}$$

Thus, the three constraints for orthogonal groups in two dimensions imply that the square of the determinant of such transformation must be equal to unity.

3. Forming the product of the the matrices corresponding to $R(\varphi_1)$ and $R(\varphi_2)$ yields

$$\begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix} \begin{pmatrix} \cos\varphi_2 & -\sin\varphi_2 \\ \sin\varphi_2 & \cos\varphi_2 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 & -\cos\varphi_1\sin\varphi_2 - \sin\varphi_1\cos\varphi_2 \\ \sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2 & -\sin\varphi_1\sin\varphi_2 + \cos\varphi_1\cos\varphi_2 \end{pmatrix}.$$

By invoking the standard trigonometric identities for the sines and cosines of the sum and difference of two angles,

$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y\,,$$
$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y\,,$$

we can write

$$\begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix} \begin{pmatrix} \cos\varphi_2 & -\sin\varphi_2 \\ \sin\varphi_2 & \cos\varphi_2 \end{pmatrix}$$

$$= \begin{bmatrix} \cos(\varphi_1+\varphi_2) & -\sin(\varphi_1+\varphi_2) \\ \sin(\varphi_1+\varphi_2) & \cos(\varphi_1+\varphi_2) \end{bmatrix}.$$

Thus,

$$R(\varphi_1+\varphi_2) = R(\varphi_1)R(\varphi_2)\,.$$

4. The expression

$$R(\varphi) = \mathrm{e}^{\varphi X}\,,$$

where

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

is defined by its Taylor series:

$$\mathrm{e}^{\varphi X} = \sum_{n=0}^{\infty} \frac{1}{n!}(\varphi X)^n\,. \tag{1}$$

Successive powers of X yield

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad X^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$X^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad X^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

whereupon this sequence is repeated. We can write this sequence in matrix form as $X^2 = -I, X^3 = -X, X^4 = I, \ldots$, where I is the 2×2 unit matrix. The powers of X are therefore given by

$$X^{2n} = \begin{cases} I, & n \text{ even} \\ -I, & n \text{ odd} \end{cases}$$

for even powers and

$$X^{2n+1} = \begin{cases} X, & n \text{ even} \\ -X, & n \text{ odd} \end{cases}$$

for odd powers Thus, the Taylor series in (1) may be written as

$$\mathrm{e}^{\varphi X} = \underbrace{\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\varphi^{2n}}_{\cos\varphi} I + \underbrace{\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\varphi^{2n+1}}_{\sin\varphi} X$$

$$= I\cos\varphi + X\sin\varphi$$

$$= \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

which is the rotation matrix in two dimensions.

5. The two parameter group

$$x' = ax + b$$

was discussed in Example 7.1. The identity was found to correspond to the parameters $a = 1$ and $b = 0$. The infinitesimal transformations are therefore given by

$$x' = (1 + \mathrm{d}a)x + \mathrm{d}b = x + x\,\mathrm{d}a + \mathrm{d}b\,.$$

If we substitute this into some function $f(x)$ and expand to first order in the parameters a and b, we obtain

$$f(x') = f(x + x\,\mathrm{d}a + \mathrm{d}b) = f(x) + x\frac{\partial f}{\partial x}\,\mathrm{d}a + \frac{\partial f}{\partial x}\,\mathrm{d}b\,,$$

from which we identify the infinitesimal operators

$$X_1 = x\frac{\partial}{\partial x}, \qquad X_2 = \frac{\partial}{\partial x}\,.$$

6. The group $C_{\infty v}$ contains all two-dimensional rotations and a vertical reflection plane, denoted by σ_v, in the x-z plane. Since this reflection changes the parity of the coordinate system, it changes the sense of the rotation angle φ. Thus, a rotation by φ in the original coordinate system corresponds to a rotation by $-\varphi$ in the

transformed coordinate system. Denoting the reflection operator by S, we then must have that

$$SR(\varphi)S^{-1} = R(-\varphi)\,. \tag{2}$$

Since $S = S^{-1}$, we can see this explicitly for the two-dimensional rotation matrix $R(\varphi)$:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix},$$

where the matrix on the right-hand side of this equation is $R(-\varphi)$. Equation (2) shows that (i) the group is no longer Abelian, and (ii) the equivalence classes correspond to rotations by φ and $-\varphi$.

7. Proceeding as in Section 7.4, the infinitesimal rotations in four dimensions which leave the quantity $x^2 + y^2 + z^2 + w^2$ invariant are

$$\begin{pmatrix} x' \\ y' \\ z' \\ w' \end{pmatrix} = \begin{pmatrix} 1 & \varphi_1 & \varphi_2 & \varphi_3 \\ -\varphi_1 & 1 & \varphi_4 & \varphi_5 \\ -\varphi_2 & -\varphi_4 & 1 & \varphi_6 \\ -\varphi_3 & -\varphi_5 & -\varphi_6 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}.$$

Substituting this coordinate transformation into a differentiable function $F(x, y, z, w)$,

$$\begin{aligned} &F(x', y', z', w') \\ &= F(x + \varphi_1 y + \varphi_2 z + \varphi_3 w, y - \varphi_1 x - \varphi_4 z + \varphi_5 w, \\ &\qquad z - \varphi_2 x + \varphi_4 y + \varphi_6 w, w - \varphi_3 x - \varphi_5 y - \varphi_6 z)\,. \end{aligned}$$

and expanding to first order in the φ_i, yields

$$F(x', y', z', w') = F(x, y, z, w) + \varphi_1\Big(y\frac{\partial F}{\partial x} - x\frac{\partial F}{\partial y}\Big)$$

$$= \varphi_2\Big(z\frac{\partial F}{\partial x} - x\frac{\partial F}{\partial z}\Big) + \varphi_3\Big(w\frac{\partial F}{\partial x} - x\frac{\partial F}{\partial w}\Big)$$
$$= \varphi_4\Big(y\frac{\partial F}{\partial z} - z\frac{\partial F}{\partial y}\Big) + \varphi_5\Big(w\frac{\partial F}{\partial y} - y\frac{\partial F}{\partial w}\Big)$$
$$= \varphi_6\Big(w\frac{\partial F}{\partial z} - z\frac{\partial F}{\partial w}\Big) .$$

From these equations and, if necessary, a change in sign of the corresponding φ_i, we can identify the following differential operators

$$A_1 = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\,, \qquad A_2 = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\,, \qquad A_3 = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\,,$$
$$B_1 = x\frac{\partial}{\partial t} - t\frac{\partial}{\partial x}\,, \qquad B_2 = y\frac{\partial}{\partial t} - t\frac{\partial}{\partial y}\,, \qquad B_3 = z\frac{\partial}{\partial t} - t\frac{\partial}{\partial z}\,.$$

8. With the infinitesimal generators calculated in Problem 7, we determine the commutators in the standard fashion. For the commutators between the A_i, we have

$$[A_1, A_2]f = A_1(A_2 f) - A_2(A_1 f)$$
$$= \Big(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\Big)\Big(x\frac{\partial f}{\partial z} - z\frac{\partial f}{\partial x}\Big) - \Big(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\Big)\Big(z\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial z}\Big)$$
$$= xz\frac{\partial^2 f}{\partial y\partial z} - z^2\frac{\partial^2 f}{\partial y\partial x} - xy\frac{\partial^2 f}{\partial z^2} + y\frac{\partial f}{\partial x} + yz\frac{\partial^2 f}{\partial z\partial x}$$
$$-x\frac{\partial f}{\partial y} - xz\frac{\partial^2 f}{\partial z\partial y} + xy\frac{\partial^2 f}{\partial z^2} + z^2\frac{\partial^2 f}{\partial x\partial y} + yz\frac{\partial^2 f}{\partial x\partial z}$$
$$= y\frac{\partial f}{\partial x} - -x\frac{\partial f}{\partial y}$$
$$= A_3 f\,,$$

$$
\begin{aligned}
[A_1, A_3]f &= A_1(A_3 f) - A_3(A_1 f) \\
&= \Big(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\Big)\Big(y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y}\Big) - \Big(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\Big)\Big(z\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial z}\Big) \\
&= z\frac{\partial f}{\partial x} + yz\frac{\partial^2 f}{\partial y \partial z} - xz\frac{\partial^2 f}{\partial y^2} - y^2\frac{\partial^2 f}{\partial z \partial x} + xy\frac{\partial^2 f}{\partial z \partial y} \\
&\quad -yz\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial x \partial z} + xz\frac{\partial^2 f}{\partial y^2} - x\frac{\partial f}{\partial z} - xy\frac{\partial^2 f}{\partial y \partial z} \\
&= z\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial z} \\
&= -A_2 f\,,
\end{aligned}
$$

$$
\begin{aligned}
[A_2, A_3]f &= A_2(A_3 f) - A_3(A_2 f) \\
&= \Big(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\Big)\Big(y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y}\Big) - \Big(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\Big)\Big(x\frac{\partial f}{\partial z} - z\frac{\partial f}{\partial x}\Big) \\
&= xy\frac{\partial^2 f}{\partial z \partial x} - x^2\frac{\partial^2 f}{\partial z \partial y} - yz\frac{\partial^2 f}{\partial x^2} + z\frac{\partial f}{\partial y} + xz\frac{\partial^2 f}{\partial x \partial y} \\
&\quad -y\frac{\partial f}{\partial z} - xy\frac{\partial^2 f}{\partial x \partial z} + yz\frac{\partial^2 f}{\partial x^2} + x^2\frac{\partial^2 f}{\partial y \partial z} - xz\frac{\partial^2 f}{\partial y \partial x} \\
&= z\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial z} \\
&= A_1 f\,.
\end{aligned}
$$

Thus, we can summarize these results as

$$[A_i, A_j] = \varepsilon_{ijk} A_k\,.$$

Similarly, for the B_i, we calculate the pertinent commutators as

$$[B_1, B_2]f = B_1(B_2 f) - B_2(B_1 f)$$

$$= \left(x\frac{\partial}{\partial w} - w\frac{\partial}{\partial x}\right)\left(y\frac{\partial f}{\partial w} - w\frac{\partial f}{\partial y}\right) - \left(y\frac{\partial}{\partial w} - w\frac{\partial}{\partial y}\right)\left(x\frac{\partial f}{\partial w} - w\frac{\partial f}{\partial x}\right)$$

$$= xy\frac{\partial^2 f}{\partial w^2} - x\frac{\partial f}{\partial y} - xw\frac{\partial^2 f}{\partial w \partial y} - yw\frac{\partial^2 f}{\partial x \partial w} + w^2\frac{\partial^2 f}{\partial x \partial y}$$

$$-xy\frac{\partial^2 f}{\partial w^2} + y\frac{\partial f}{\partial x} + yw\frac{\partial^2 f}{\partial w \partial x} + xw\frac{\partial^2 f}{\partial y \partial t} - w^2\frac{\partial^2 f}{\partial y \partial x}$$

$$= y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y}$$

$$= A_3 f\,,$$

$$[B_1, B_3]f = B_1(B_3 f) - B_3(B_1 f)$$

$$= \left(x\frac{\partial}{\partial w} - w\frac{\partial}{\partial x}\right)\left(z\frac{\partial f}{\partial w} - w\frac{\partial f}{\partial z}\right) - \left(z\frac{\partial}{\partial w} - w\frac{\partial}{\partial z}\right)\left(x\frac{\partial f}{\partial w} - w\frac{\partial f}{\partial x}\right)$$

$$= xz\frac{\partial^2 f}{\partial w^2} - x\frac{\partial f}{\partial z} - xw\frac{\partial^2 f}{\partial w \partial z} - zw\frac{\partial^2 f}{\partial x \partial w} + t^2\frac{\partial^2 f}{\partial x \partial z}$$

$$-xz\frac{\partial^2 f}{\partial w^2} + z\frac{\partial f}{\partial x} + zw\frac{\partial^2 f}{\partial w \partial x} + xt\frac{\partial^2 f}{\partial z \partial w} - w^2\frac{\partial^2 f}{\partial z \partial x}$$

$$= z\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial z}$$

$$= -A_2 f\,,$$

$$[B_2, B_3]f = B_2(B_3 f) - B_3(B_2 f)$$

$$= \left(y\frac{\partial}{\partial w} - w\frac{\partial}{\partial y}\right)\left(z\frac{\partial f}{\partial w} - w\frac{\partial f}{\partial z}\right) - \left(z\frac{\partial}{\partial w} - w\frac{\partial}{\partial z}\right)\left(y\frac{\partial f}{\partial w} - w\frac{\partial f}{\partial y}\right)$$

$$= yz\frac{\partial^2 f}{\partial w^2} - y\frac{\partial f}{\partial z} - yw\frac{\partial^2 f}{\partial w \partial z} - zw\frac{\partial^2 f}{\partial y \partial w} + w^2\frac{\partial^2 f}{\partial y \partial z}$$

$$-yz\frac{\partial^2 f}{\partial w^2}+z\frac{\partial f}{\partial y}+zw\frac{\partial^2 f}{\partial w\partial y}+yt\frac{\partial^2 f}{\partial z\partial w}-w^2\frac{\partial^2 f}{\partial z\partial y}$$

$$=z\frac{\partial f}{\partial y}-y\frac{\partial f}{\partial z}$$

$$=A_1 f\,.$$

These results can be summarized as

$$[B_i,B_j]=\varepsilon_{ijk}A_k\,.$$

Finally, for the commutators between the A_i and B_j, we note first by inspection that

$$[A_i,B_i]=0\,,$$

for $i=1,2,3$, since A_i and B_i involve mutually exclusive pairs of variables. For the remaining commutator pairs, we have

$$[A_1,B_2]=A_1(B_2f)-B_2(A_1f)$$

$$=\Big(z\frac{\partial}{\partial y}-y\frac{\partial}{\partial z}\Big)\Big(y\frac{\partial f}{\partial w}-w\frac{\partial f}{\partial y}\Big)-\Big(y\frac{\partial}{\partial w}-w\frac{\partial}{\partial y}\Big)\Big(z\frac{\partial f}{\partial y}-y\frac{\partial f}{\partial z}\Big)$$

$$=z\frac{\partial f}{\partial w}+yz\frac{\partial^2 f}{\partial y\partial w}-zw\frac{\partial^2 f}{\partial y^2}-y^2\frac{\partial^2 f}{\partial z\partial w}+yw\frac{\partial^2 f}{\partial z\partial y}$$

$$-yz\frac{\partial^2 f}{\partial w\partial y}+y^2\frac{\partial^2 f}{\partial w\partial z}+zw\frac{\partial^2 f}{\partial y^2}-w\frac{\partial f}{\partial z}-yw\frac{\partial^2 f}{\partial y\partial z}$$

$$=z\frac{\partial f}{\partial w}-w\frac{\partial f}{\partial z}$$

$$=B_3 f\,,$$

$$[A_1,B_3]=A_1(B_3f)-B_3(A_1f)$$

$$=\Big(z\frac{\partial}{\partial y}-y\frac{\partial}{\partial z}\Big)\Big(z\frac{\partial f}{\partial w}-w\frac{\partial f}{\partial z}\Big)-\Big(z\frac{\partial}{\partial w}-w\frac{\partial}{\partial z}\Big)\Big(z\frac{\partial f}{\partial y}-y\frac{\partial f}{\partial z}\Big)$$

$$= z^2\frac{\partial^2 f}{\partial y\partial w} - zw\frac{\partial^2 f}{\partial y\partial z} - y\frac{\partial f}{\partial w} - yz\frac{\partial^2 f}{\partial z\partial w} + yw\frac{\partial^2}{\partial z^2}$$

$$-z^2\frac{\partial^2 f}{\partial w\partial y} + yz\frac{\partial^2 f}{\partial w\partial z} + w\frac{\partial f}{\partial y} + zw\frac{\partial^2 f}{\partial z\partial y} - yw\frac{\partial^2 f}{\partial z^2}$$

$$= w\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial w}$$

$$= -B_2 f\,,$$

$$[A_2, B_3] = A_2(B_3 f) - B_3(A_2 f)$$

$$= \left(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\right)\left(z\frac{\partial f}{\partial w} - w\frac{\partial f}{\partial z}\right) - \left(z\frac{\partial}{\partial w} - w\frac{\partial}{\partial z}\right)\left(x\frac{\partial f}{\partial z} - z\frac{\partial f}{\partial x}\right)$$

$$= x\frac{\partial f}{\partial w} + xz\frac{\partial^2 f}{\partial z\partial w} - xw\frac{\partial^2 f}{\partial z^2} - z^2\frac{\partial^2 f}{\partial x\partial w} + zw\frac{\partial^2 f}{\partial x\partial z}$$

$$-xz\frac{\partial^2 f}{\partial w\partial z} + z^2\frac{\partial^2 f}{\partial w\partial x} - w\frac{\partial f}{\partial x} + xw\frac{\partial^2 f}{\partial z^2} - zw\frac{\partial^2 f}{\partial x\partial z}$$

$$= x\frac{\partial f}{\partial w} - w\frac{\partial f}{\partial x}$$

$$= B_1 f\,.$$

Thus,

$$[A_i, B_j] = \varepsilon_{ijk} B_k\,.$$

9. Consider the following linear combinations of the operators in Problem 7:

$$J_i = \tfrac{1}{2}(A_i + B_i), \qquad K_i = \tfrac{1}{2}(A_i - B_i)\,. \tag{3}$$

We can now use the commutation relations derived in Problem 8 to derive the commutation relations for the J_i and K_j. For the J_i, we have

$$\begin{aligned}
[J_i, J_j] &= \tfrac{1}{4}[A_i + B_i, A_j + B_j] \\
&= \tfrac{1}{4}\Big([A_i, A_j] + [A_i, B_j] + [B_i, A_j] + [B_i, B_j]\Big) \\
&= \tfrac{1}{4}(\varepsilon_{ijk}A_k + \varepsilon_{ijk}B_k + \varepsilon_{ijk}B_k + \varepsilon_{ijk}A_k) \\
&= \varepsilon_{ijk}\tfrac{1}{2}(A_k + B_k) \\
&= \varepsilon_{ijk}J_k\,,
\end{aligned}$$

$$\begin{aligned}
[K_i, K_j] &= \tfrac{1}{4}[A_i - B_i, A_j - B_j] \\
&= \tfrac{1}{4}\Big([A_i, A_j] - [A_i, B_j] - [B_i, A_j] + [B_i, B_j]\Big) \\
&= \tfrac{1}{4}(\varepsilon_{ijk}A_k - \varepsilon_{ijk}B_k - \varepsilon_{ijk}B_k + \varepsilon_{ijk}A_k) \\
&= \varepsilon_{ijk}\tfrac{1}{2}(A_k - B_k) \\
&= \varepsilon_{ijk}K_k\,,
\end{aligned}$$

$$\begin{aligned}
[J_i, K_j] &= \tfrac{1}{4}[A_i + B_i, A_j - B_j] \\
&= \tfrac{1}{4}\Big([A_i, A_j] - [A_i, B_j] + [B_i, A_j] - [B_i, B_j]\Big) \\
&= \tfrac{1}{4}(\varepsilon_{ijk}A_k + \varepsilon_{ijk}B_k - \varepsilon_{ijk}B_k - \varepsilon_{ijk}A_k) \\
&= 0\,.
\end{aligned}$$

Group Theory

Problem Set 10 December 11, 2001

Note: Problems marked with an asterisk are for Rapid Feedback.

1.* Prove that a proper orthogonal transformation in an odd-dimensional space always possesses an axis, i.e., a line whose point are left unchanged by the transformation.

2. Prove **Euler's theorem**: The general displacement of a rigid body with one fixed point is a rotation about an axis.

3.* The functions $(x \pm iy)^m$, where m is an integer generate irreducible representations of SO(2). Suppose we now consider the group O(2), where we now allow *improper* rotations. Use Schur's lemma to show that these functions generate irreducible *two*-dimensional representations of O(2) for $m \neq 0$, but a *reducible* representation for $m = 0$.

Hint: The general improper rotation in two dimensions is

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

4. Consider the rotation matrix obtained by rotating an initial set of axes counterclockwise by ϕ about the z-axis, then rotated about the new x-axis counterclockwise by θ, and finally rotated about the new z-axis counterclockwise by ψ. These are the **Euler angles** and the corresponding rotation matrix is

$$\begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix}$$

Verify that the angle of rotation φ of this transformation is given by

$$\cos\left(\tfrac{1}{2}\varphi\right) = \cos\left[\tfrac{1}{2}(\phi + \psi)\right]\cos\left(\tfrac{1}{2}\theta\right)$$

5. Determine the axis of the transformation in Problem 4.

6.* Verify that the direct product of two irreducible representations of SO(3) has the following decomposition

$$\chi^{(\ell_1)}(\varphi)\chi^{(\ell_2)}(\varphi) = \sum_{\ell=|\ell_1-\ell_2|}^{\ell_1+\ell_2} \chi^{(\ell)}(\varphi)$$

This is called the **Clebsch–Gordan series** and provides a group-theoretic statement of the addition of angular momenta.

7.* Determine the corresponding Clebsch-Gordan series for SO(2).

8.* Show that the requirement that xx^*+yy^* is invariant under the complex transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

together with the determinant of this transformation being unity means that the transformation must have the form

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

where $aa^* + bb^* = 1$.

Group Theory

Solutions to Problem Set 10 December 14, 2001

1. As shown in Section 8.3.1, the eigenvalues of an orthogonal matrix have modulus unity. These eigenvalues are also the roots of the polynomial equation $\det(A - \lambda I) = 0$, so the Fundamental Theorem of Algebra requires that, if these roots are complex, they must occur in complex conjugate pairs. Thus, only in an *odd-dimensional* space is there guaranteed to be a single real eigenvalue of unity. The corresponding eigenvector is the axis of rotation.

2. If the fixed point is taken as the origin of the set of axes of the body, then the displacement of the rigid body involves no translation, but only a change of orientation, i.e., a rotation. Since, in three dimensions, every rotation can be expressed in an axis-angle representation, Euler's theorem follows immediately.

3. The general improper transformation in two dimensions is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} .$$

Thus, for the functions $(x \pm iy)^m$ we have

$$\begin{aligned} (x' \pm iy')^m &= \Big[x\cos\varphi + y\sin\varphi \pm i(x\sin\varphi - y\cos\varphi)\Big]^m \\ &= \Big[x(\cos\varphi \pm i\sin\varphi) \mp iy(\cos\varphi \pm i\sin\varphi)\Big]^m \\ &= (x \mp iy)^m \, \mathrm{e}^{\pm im\varphi} \, , \end{aligned}$$

so they generate the representation

$$\begin{bmatrix} (x' + iy')^m \\ (x' - iy')^m \end{bmatrix} = \begin{pmatrix} 0 & \mathrm{e}^{im\varphi} \\ \mathrm{e}^{-im\varphi} & 0 \end{pmatrix} \begin{bmatrix} (x + iy)^m \\ (x - iy)^m \end{bmatrix} .$$

To determine whether or not this representation is reducible, we apply Schur's first lemma. Suppose a matrix A commutes with all of the matrices of our two dimensional representation. Then, we have

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 & \mathrm{e}^{im\varphi} \\ \mathrm{e}^{-im\varphi} & 0 \end{pmatrix}}_{\begin{pmatrix} a_{12}\mathrm{e}^{-im\varphi} & a_{11}\mathrm{e}^{im\varphi} \\ a_{22}\mathrm{e}^{-im\varphi} & a_{21}\mathrm{e}^{im\varphi} \end{pmatrix}} = \underbrace{\begin{pmatrix} 0 & \mathrm{e}^{im\varphi} \\ \mathrm{e}^{-im\varphi} & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}_{\begin{pmatrix} a_{21}\mathrm{e}^{im\varphi} & a_{22}\mathrm{e}^{im\varphi} \\ a_{11}\mathrm{e}^{-im\varphi} & a_{12}\mathrm{e}^{-im\varphi} \end{pmatrix}} .$$

Thus, if $m \neq 0$, we must require that $a_{12} = a_{21} = 0$ and that $a_{11} = a_{22}$, i.e., A is multiple of the 2×2 unit matrix and, according to Schur's first lemma, this representation is *irreducible*. However, of $m = 0$, we need only require that $a_{12} = a_{21}$ and $a_{11} = a_{22}$, so this is a *reducible* representation.

4. The rotation angle φ is calculated from the trace of the transformation matrix:

$$\begin{aligned} 1 + 2\cos\varphi &= \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi - \sin\psi\sin\phi \\ &\qquad + \cos\theta\cos\phi\cos\psi + \cos\theta \\ &= (1+\cos\theta)(\cos\phi\cos\psi - \sin\phi\sin\psi) + \cos\theta \\ &= (1+\cos\theta)\cos(\phi+\psi) + \cos\theta\,. \end{aligned}$$

Using the triginometric identity

$$1 + 2\cos\varphi = 4\cos^2\left(\tfrac{1}{2}\varphi\right) - 1\,,$$

we obtain

$$\begin{aligned} 4\cos^2\left(\tfrac{1}{2}\varphi\right) &= (1+\cos\theta)[1+\cos(\phi+\psi)] \\ &= 4\cos^2\left(\tfrac{1}{2}\theta\right)\cos^2\left[\tfrac{1}{2}(\phi+\psi)\right], \end{aligned}$$

or,

$$\cos\left(\tfrac{1}{2}\varphi\right) = \cos\left(\tfrac{1}{2}\theta\right)\cos\left[\tfrac{1}{2}(\phi+\psi)\right].$$

5. The axis of the transformation in Problem 4 is determined from the equations derived in Section 8.3.2:

$$\frac{n_2}{n_1} = \frac{a_{31}-a_{13}}{a_{23}-a_{32}}, \qquad \frac{n_3}{n_1} = \frac{a_{12}-a_{21}}{a_{23}-a_{32}}.$$

The denominator of these expressions is

$$a_{23}-a_{32} = \sin\theta\cos\psi + \sin\theta\cos\phi = \sin\theta(\cos\psi+\cos\phi).$$

We also have

$$\begin{aligned} a_{31}-a_{13} &= \sin\theta\sin\phi - \sin\theta\sin\psi = \sin\theta(\sin\phi-\sin\psi) \\ a_{12}-a_{21} &= \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi \\ &\qquad + \sin\psi\cos\phi + \cos\theta\sin\phi\cos\psi \\ &= (1+\cos\theta)(\cos\phi\sin\psi + \sin\phi\cos\psi) \\ &= (1+\cos\theta)\sin(\phi+\psi). \end{aligned}$$

Thus, the (unnormalized) direction of the rotation axis is

$$\left\{1,\ \frac{\sin\phi-\sin\psi}{\cos\psi+\cos\phi},\ 2\frac{(1+\cos\theta)\sin(\phi+\psi)}{\sin\theta(\cos\phi+\cos\psi)}\right\}.$$

6. There are a number of ways of decomposing the direct product of irreducible representations of SO(3). The books by Tinkham

and Jones give two very different approaches. Below, we provide a third method. We first calculate the direct product

$$\chi^{(\ell)}(\varphi)\chi^{(1)}(\varphi) = \left(\sum_{m=-\ell}^{\ell} \mathrm{e}^{-im\varphi}\right)\left(\sum_{m_1=-1}^{1} \mathrm{e}^{-im_1\varphi}\right).$$

By expanding the second summation and multiplying the first summation with each of the exponentials, we obtain

$$\begin{aligned}\left(\sum_{m=-\ell}^{\ell} \mathrm{e}^{-im\varphi}\right)\left(\sum_{m_1=-1}^{1} \mathrm{e}^{-im_1\varphi}\right) &= \sum_{m=-\ell}^{\ell} \mathrm{e}^{-im\varphi}\left(\mathrm{e}^{i\varphi} + 1 + \mathrm{e}^{-i\varphi}\right) \\ &= \sum_{m=-\ell}^{\ell} \mathrm{e}^{-i(m-1)\varphi} + \sum_{m=-\ell}^{\ell} \mathrm{e}^{-im\varphi} + \sum_{m=-\ell}^{\ell} \mathrm{e}^{-i(m+1)\varphi}.\end{aligned}$$

If, in the first summation on the right-hand side of this equation, we change the summation variable to $m' = m - 1$ and in the last summation change the summation variable to $m' = m + 1$, we have

$$\begin{aligned}&\sum_{m=-\ell}^{\ell} \mathrm{e}^{-i(m-1)\varphi} + \sum_{m=-\ell}^{\ell} \mathrm{e}^{-i(m+1)\varphi} \\ &\qquad = \sum_{m'=-\ell-1}^{\ell-1} \mathrm{e}^{-im'\varphi} + \sum_{m'=-\ell+1}^{\ell+1} \mathrm{e}^{-im'\varphi} \\ &\qquad = \sum_{m'=-(\ell+1)}^{\ell+1} \mathrm{e}^{-im'\varphi} + \sum_{m'=-(\ell-1)}^{\ell-1} \mathrm{e}^{-im'\varphi}.\end{aligned}$$

In fact, for any positive integer k, we have

$$\begin{aligned}&\sum_{m=-\ell}^{\ell} \mathrm{e}^{-i(m-k)\varphi} + \sum_{m=-\ell}^{\ell} \mathrm{e}^{-i(m+k)\varphi} \\ &\qquad = \sum_{m'=-\ell-k}^{\ell-k} \mathrm{e}^{-im'\varphi} + \sum_{m'=-\ell+k}^{\ell+k} \mathrm{e}^{-im'\varphi} \\ &\qquad = \sum_{m'=-(\ell+k)}^{\ell+k} \mathrm{e}^{-im'\varphi} + \sum_{m'=-(\ell-k)}^{\ell-k} \mathrm{e}^{-im'\varphi}. \qquad (1)\end{aligned}$$

Thus, we conclude that

$$\chi^{(\ell)}(\varphi)\chi^{(1)}(\varphi) = \chi^{(\ell-1)}(\varphi) + \chi^{(\ell)}(\varphi) + \chi^{(\ell+1)}(\varphi)\,.$$

Then, by using (1), we have, in the general case

$$\begin{aligned}\chi^{(\ell)}(\varphi)\chi^{(\ell')}(\varphi) &= \sum_{m=-\ell}^{\ell} \mathrm{e}^{-im\varphi}\Big[\mathrm{e}^{i\ell'\varphi} + \mathrm{e}^{i(\ell'-1)\varphi} + \cdots + \mathrm{e}^{-i\ell'\varphi}\Big] \\ &= \chi^{(\ell+\ell')}(\varphi) + \chi^{(\ell+\ell'-1)}(\varphi) + \cdots + \chi^{(\ell-\ell')}\,.\end{aligned}$$

Therefore,

$$\chi^{(\ell)}(\varphi)\chi^{\ell'}(\varphi) = \sum_{m=\ell-\ell'}^{\ell+\ell'} \chi^{(m)}(\varphi)\,,$$

where, from our procedure, it is clear that $\ell' \leq \ell$.

7. The corresponding Clebsch–Gordan series for SO(2) is very simple because the group is Abelian. Since

$$\chi^{(m)}(\varphi) = \mathrm{e}^{im\varphi}\,,$$

then

$$\chi^{(m_1)}(\varphi)\chi^{(m_2)}(\varphi) = \chi^{(m_1+m_2)}(\varphi)\,.$$

8. Given the complex transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\,, \tag{2}$$

then the invariance of the quantity $xx^* + yy^*$ yields

$$\begin{aligned} &x'x'^* + y'y'^* \\ &= (ax + by)(a^*x^* + b^*y^*) + (cx + dy)(c^*x^* + d^*y^*) \\ &= (aa^* + cc^*)xx^* + (ab^* + cd^*)xy^* + (a^*b + c^*d)x^*y \\ &+ (cc^* + dd^*)yy^* \,. \end{aligned}$$

Maintaining equality for all independent variations of x and y requires that

$$aa^* + cc^* = 1, \qquad ab^* + cd^* = 0, \qquad cc^* + dd^* = 1\,. \tag{3}$$

A fourth condition is that the determinant of the transformation in (2) is unity:

$$ad - bc = 1 \tag{4}$$

If we take the second of equations (3), multiply by a^*, and then use the first of these equations and Equation (4), we obtain

$$\begin{aligned} a^*(ab^* + cd^*) &= aa^*b^* + a^*cd^* \\ &= (1 - cc^*)b^* + (1 + b^*c^*)c \\ &= b^* + c = 0 \end{aligned}$$

which yields

$$c = -b^*$$

The second of equations (4) then immediately yields

$$a = d^* \tag{5}$$

Thus, the transformation (2) must have the form

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

where $aa^* + bb^* = 1$.